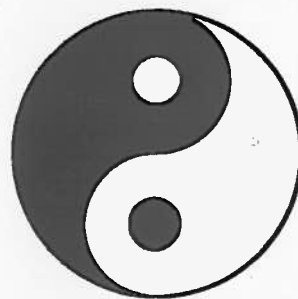


Université Catholique de Louvain

Faculté des Sciences

**Morphologie des aires de marché
dans le plan muni de la distance euclidienne.**

Le modèle de gravitation généralisé



**Dissertation présentée en vue de
l'obtention du grade de
Docteur en Sciences
par Pierre HANJOUL**

**Promoteur : Jacques THISSE
Copromoteur : Dominique PEETERS**

Louvain-la-Neuve

7 mai 1996

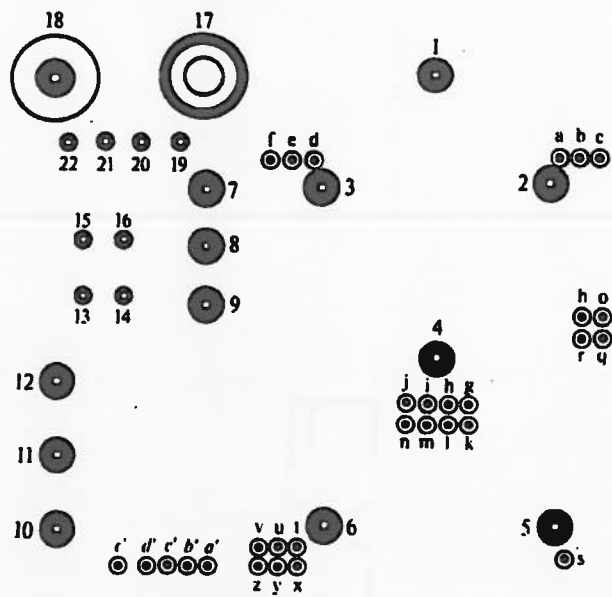
A mon père et à ma soeur,
à ma famille, celle de cœur aussi;
à la communauté de la Poudrière,
à mes amis de l'unité GEOG et de l'ILV,
à mes promoteur, copromoteur, lecteurs,
à tous ceux et celles, vivants et disparus,
qui, ici et ailleurs, maintenant et autrefois,
me soutiennent de leur affection.
Je ne pourrais les citer tous,
mais qu'ils sachent bien
que je pense fort à eux en ce moment
et que certains, non nommés, ne sont pas ici les moindres.

Je les remercie de cet appui irremplaçable,
je demande aussi leur indulgence.
Qu'ils veuillent bien me pardonner
de les avoir négligés ces derniers temps
pour terminer ceci.
Puissent-ils m'aider à ce que cette étape de ma vie
ne produise pas seulement des fruits dérisoires.

Ce n'est pas par jeu que Nous avons créé les cieux
et la Terre et ce qui est entre les deux
(Coran)

Quand je parlerais les langues des hommes et des anges,
si je n'ai pas la charité, je ne suis qu'un gong retentissant,
qu'une cymbale tonitruante.
Que j'aie l'inspiration, que je sache tous les mystères
et toute la connaissance,
que j'aie la foi à transporter les montagnes,
si je n'ai pas la charité,
je ne suis rien.
(Saint Paul)

Ma petite âme s'en est allée avec la veille;
il faut maintenant dire des choses nouvelles.
Les mots, si nombreux soient-ils, appartiennent à la veille;
il faut maintenant dire des choses nouvelles.
(Mevlânâ Jalâl ad-Dîn Rûmî)



Introduction

Part I. Two centres

Chapter 1. Modelling consumer behaviour

1.1. All-or-nothing (i.e., nondistributive) interaction models	1.1
1.2. The generalized gravity model.....	1.3
1.3. Constrained generalized gravity models.....	1.7
1.4. Models attempting to justify the generalized gravity model	1.9
1.5. Other models reducible to GG or all-or-nothing models	1.10
1.6. The multinomial logit model.....	1.12
1.7. Derived issues.....	1.14

Chapter 2. A few introductory properties

2.1. Delimitation of the main problem and definitions.....	2.1
2.2. The identification problem.....	2.3
2.3. Position.....	2.7
2.4. Dependence on Q and δ_{jk}	2.8
2.5. Some easy cases.....	2.10

**Chapter 3. Variation of the transportation cost difference
(and of y) along specific curves**

3.1. Move along circles centred on a point of $[jk]$	3.1
3.2. Radial moves from points of $[jk]$	3.2
3.3. Move along horizontal straight lines.....	3.5
3.4. Crossing between indifference lines and other hypercircles	3.6
3.5. Relative position of the indifference line and some curves associated with another t.c.f.	3.9
3.6. On parallel association.....	3.14
3.7. Variation of y with x along the indifference line.....	3.18

Table of contents .2.

Chapter 4. Remarkable bounds and asymptotic behaviour: spatial extension

4.1. Emptiness of market area Z_k	4.3
4.2. Boundedness of market area Z_k	4.4
4.3. Position of centres relatively to market areas.....	4.8
4.4. Circular bounds.....	4.9
4.5. Hyperbolic bounds.....	4.16
4.6. Asymptotic direction.....	4.18
4.7. Vertical bounds and vertical limiting lines.....	4.19
4.8. Oblique limiting lines.....	4.21
4.9. Horizontal limiting lines.....	4.24

Chapter 5. The standpoint of mathematical topology: spatial structure

5.1. Arcwise-connectedness.....	5.1
5.2. Simple connectedness.....	5.5
5.3. Spinedness.....	5.7
5.4. Starshapedness and visibility.....	5.8
5.5. Convexity.....	5.12
5.6. About interiors, boundaries, and closures.....	5.16
5.7. Zero-dimensional market areas.....	5.18

Part II. Several centres

Chapter 6. Generalizing the concept of market area

6.1. Individual comparison with other centres.....	6.1
6.2. Global comparison with other centres.....	6.7
6.3. Other definitions.....	6.11

Chapter 7. Starshapedness and position

7.1. Starshapedness.....	7.1
7.2. Position.....	7.11
7.3. About some finite-range distance deterrence functions.....	7.13
7.4. Choice, share, median, and average areas.....	7.17

Table of contents .3.

Chapter 8. Convexity and arcwise-connectedness

- 8.1. Convexity 8.1
- 8.2. The three-centre problem 8.4
- 8.3. Arcwise-connectedness 8.12
- 8.4. Choice and preference areas 8.19

Chapter 9. Other theoretical results

- 9.1. Approximations of dominance areas 9.1
- 9.2. Total potential 9.9

Chapter 10. Confrontation with reality..

- 10.1. Zero-dimensional areas 10.1
- 10.2. Market and shared influence area 10.2
- 10.3. Share areas 10.6

Conclusion

Appendix: a program written with Mathematica

Index

References

Illustrations hors thèse (explication)

Thèse annexe :

Lorsque les coûts de transport sont une fonction concave de la distance, dans le problème d'Hotelling (1929), il n'y a pas d'équilibre parfait symétrique en prix et en localisation tel que dans un voisinage des localisations d'équilibre existe un seul équilibre en prix variant continûment avec la localisation des firmes.

Introduction

In some immemorial era, when I was being employed by the department of Geography in Louvain-la-Neuve, I was asked by Hubert Beguin and Jean-Claude Thill to read a seven-page draft of a paper they were writing about the shape assumed by the market areas of two centres when those areas are generated according to the generalized gravity model. I found their idea interesting but also expressed that I had the feeling that the subject could be developed. Hubert answered in his famous right-to-the-point manner: 'Well then , develop it', or something similar, and it became my job for a period of three years, during which I frequently discussed its developments with Hubert and Jean-Claude. The result was published in 1988 under the title *Theoretical Market Areas under Euclidean Distance*, with me, Hubert and Jean-Claude as co-authors. We also wrote several papers in its perspective [Hanjoul et al. (1986, 1987, 1989), Hanjoul and Thill (1987)].

Subsequently, as I had become by far the main author, they suggested me to transform it into a doctoral thesis; Jacques Thisse accepted to be my director. The only (seemingly innocuous) requirement of the Department of Geography was that I should add something to make it more applicable. It became rapidly clear to me that the best way to do this was to extend the reflection to situations involving any number of centres and to compare it with existing economic maps. However, it appeared that the extended problem was not easy and it took me new years of research, the more since I was not employed any more in that department and I was absolutely no more able nor willing to dedicate all my energy to it.

At the end of 1994 I was in possession of all my main results and thinking in terms of writing everything accurately down. In particular, I began to rewrite completely our work of 1988. But this was too ambitious, and after a few chapters I saw I had better to select the chapters I should keep for the thesis. Also, Dominique Peeters, who had meanwhile become my co-director, asked me to bring my research to an end because he was planning a long stay abroad. He so succeeded where the kind and endless supplications of Hubert, and the more sparse and resigned ones of Henry Zoller, had failed so far.

The two-centre problem forms the five chapters of the first part of the thesis. The first chapter, rather a new one, is an introduction to the problem and its historical developments. The four other chapters have been taken from the work of 1988; they concern some introductory results (Chapter 2), basic results to be used in the subsequent chapters (Chapter 3), geometrical bounds on the market areas (Chapter 4), and topological properties (Chapter 5). I have selected them either because I wanted to rewrite some sections of them, making a kind of new edition (this is particularly true for Chapter 3), or because, even though not related with the second part of the thesis, they gave the opportunity to show that market areas may have quite unexpected shapes (e.g. Section 2.1 and Chapter 4) ¹. I am conscious, however, that, for clarity's sake, it would have been preferable to drop some results. And also, probably, not to make Chapter 3 a separate one; but time was missing to perform such a deep change.

The second part of the thesis is completely new. It investigates the possible ways to extend the notion of market area from a situation with two centres to one with several centres (Chapter 6). The main and most general property, starshapedness, is treated in Chapter 7; while two others, convexity and arcwise-connectedness, are grouped in Chapter 8. Chapter 9 contains a few additional results about possible approximations of one of the extensions of the notion of market area, and about total potential. Chapter 10 makes a comparison with existing economic maps.

A last word before beginning: notice that we shall not investigate equilibrium issues. The distribution of centres over the plane is given, and so are the prices or attractivenesses. We do not ask, similarly, whether the market area is large enough to ensure profit; such a question would be a nonsense here, since we do not know the customers' locations; as will be seen now.

¹ The chapters of Hanjoul et al. (1988) that I have left aside concern: the measurement of market areas (Chapter 9), dependence of market areas on the exponent a when the transportation costs are proportional to δ^a , Euclidean distance being denoted as δ (Chapter 10), properties of two limiting shapes of market areas (Chapter 11), effects of a relaxation of some of our assumptions on the relation of transportation costs to distance (Chapter 12), and Cartesian ovals (Chapter 13).

Part I. Two centres

1. Modelling consumer behaviour

1.1. All-or-nothing (i.e., nondistributive) interaction models

Consider a set of urban centres, C , and any place i of the geographical space. We assume that some commodity is sold in the centres at different given mill prices; the price in centre c is p_c . The distance being denoted by δ , we assume transportation costs to be simply a function of distance, $h(\delta)$. The delivered price for buying the commodity in c is thus $p_c + h(\delta_{ic})$. The consumer located in i , if we assume it to behave rationally [as in the theory developed by Christaller (1933), for instance], will thus buy the commodity in the centre c when $p_c + h(\delta_{ic})$ is minimal.

When C is reduced to a two-centre set, $\{j, k\}$, the interaction t_{ik} of the consumer place i with a centre k may thus be described as follows :

$$\begin{aligned} t_{ik} &= \mathcal{D}_i & \text{if } h(\delta_{ij}) - h(\delta_{ik}) > Q, \\ t_{ik} &= 0 & \text{if } h(\delta_{ij}) - h(\delta_{ik}) < Q, \\ t_{ik} &= \mathcal{D}_i / 2 & \text{if } h(\delta_{ij}) - h(\delta_{ik}) = Q, \end{aligned}$$

where

$$Q = p_k - p_j,$$

$\mathcal{D}_i$ being the total demand in i . In the third case, when $h(\delta_{ij}) - h(\delta_{ik}) = Q$, the consumer makes no difference between j and k , and it is generally assumed that it buys equally in j and k .

The area defined by the inequality

$$h(\delta_{ij}) - h(\delta_{ik}) \geq Q \tag{1.1}$$

is thus the set of points where buying in k is not less advantageous than in j . It is generally known as the *market area* (*natural area*) of k ; we denote it by $Z_k(Q)$, or simply by Z_k ; Z_j is similarly defined. We shall see in Chapter 6 how to extend the definition to the case of several centres. The set of points where the delivered prices relative to both centres are equal (i.e., $Z_j \cap Z_k$) has received the name of *isostant* from Schilling (1924, 1925); the points of $Z_j \cap Z_k$ have been known as *breaking points* since Strohkarck and Phelps (1948), though not in the perspective of all-or-nothing models.

The shape of $Z_j \cap Z_k$ when the transportation space is a plane and the transportation cost $h(\delta)$ is the Euclidean distance δ itself seems to have first been studied by Rau (1841); it is a branch of a hyperbola with foci j and k . Its equation: $|\delta_{ij} - \delta_{ik}| = Q$, known since Fetter (1924) as the *Economic Law of market areas*, has in fact been known to describe such a hyperbola since the Greek mathematician Apollonius of Perga (†180 BC), but the rediscoveries and developments of that result form a long story studied by Shieh (1985), Robine (1985), and partially by Ponsard (1983) and Ekelund and Shieh (1986). After Rau, the main names of that story are those of Launhardt (1882), Cheysson (1887), Fetter (op. cit.), Palander (1935), Schneider (1935), and Hyson and Hyson (1950).

More generally, we may assume a general additive utility model and express as follows the utility u_{ij} , for a possible consumer located in a place i , of interacting with a centre j :

$$u_{ij} = - h(\delta_{ij}) - p_j,$$

where p and h have a no defined physical significance: $- p_j$ is a measure of the attraction exerted by centre j , and $h(\delta)$ is some increasing function of distance. The consumer chooses to interact with centre k rather than with centre j when $u_{ik} > u_{ij}$, and vice versa. A place i belongs to the market area Z_k of centre k iff. $u_{ik} \geq u_{ij}$, which condition is similar to (1.1). That simple model forms the basis of more elaborate ones, some of which appear in the following sections of the present chapter.

Of course, we also obtain the formulation (1.1) if we use a multiplicative utility model. Defining indeed utility as

$$\acute{u}_{ij} = a_j F(\delta_{ij}),$$

where a_j is another measure of the attraction of j and F is a positive and decreasing function of distance, we just have to set $u = \ln(\acute{u})$, 'ln' standing for the Napierian logarithm, and

$$p_j = - \ln(a_j)$$

$$h = - \ln \circ F$$

to be back to (1.1).

On the other hand, Rau and his successors, in studying the basic additive all-or-nothing model, generally allowed their transportation rate r to depend on the centre with which interaction occurs; for them, the transportation cost to centre c is equal to $r_c \delta_{ic}$. Otherwise said, they consider different *accessibilities* to the centres and do not integrate them into the attraction coefficients $-p_c$. When $C = \{j, k\}$, the inequality describing Z_k is thus

$$r_j \delta_{ij} - r_k \delta_{ik} \geq p_k - p_j .$$

The reader is referred to Hanjoul *et al.*, Niérat (1987), and Parr (1995a) for a study of the boundary of that area. It is known to be a particular quartic curve, the *Cartesian oval*, named so in honour of Descartes (1637) who discovered them as a solution to a problem in optics. That curve is reduced to a circle when the right-hand side of (1.2) is zero (i.e., if $p_k = p_j$), a property also known to Apollonius of Perga.

A generalized form of that model would be

$$r_j g(\delta_{ij}) - r_k g(\delta_{ik}) \geq p_k - p_j , \quad (1.2)$$

where the coefficients r_c are positive and g is some increasing function. If the function g is quadratic, the boundary, obviously, is a circle. As to the general case, Hanjoul *et al.* have shown in their Section 2.2 the partial reducibility of (1.2) to (1.1).

1.2. The generalized gravity model

Another model used to represent the interaction between places is the *generalized gravity model* or *spatial interaction model*; see e.g. Sheppard (1978):

$$t_{ic} = m_i a_c F(\delta_{ic}) , \quad (1.3)$$

where t_{ic} is the amount of *interaction* between a centre c and a place i , due to the relative influence exerted by c on i ; t_{ic} corresponds to the concept of *energy* introduced by Stewart (1948), using an obvious physical analogy ¹.

¹ A similar formula was proposed by Zipf (1946). Reilly (1929, 1931) was the first important author using the gravity hypothesis in human sciences, but he only computed interaction ratios between cities. Carey (1858-1859) already formulated a hypothesis similar to (1.3).

'F', 'm', and 'a' have (*a priori*) another meaning than in Section 1.1. The coefficient a_c , still measuring the attraction of c , is the *attractiveness* ('*attractivité*') of the centre c ; m_i , the *mass* of i , or its *emissivity* ('*émissivité*'), is a function of the attributes of the place i . $F(\delta_{ic})$ is a decreasing and strictly positive function of the distance δ_{ic} between i and c ; it is also continuous everywhere, but *lato sensu* in 0 since we admit the possibility that $F(0) = \infty$. That function is called the function of *distance deterrence*, *distance decay* or *distance friction*.

In fact, Sheppard's formula contains a proportionality factor, K : the right-hand side of his equality would write $K m_i a_c F(\delta_{ic})$. We include K in the other factors by redimensioning them. In particular, defining the mass m_i as the total demand in i , D_i , is not always convenient since $F(0)$ may be infinite; in such a case, we may assume for instance that $F(1) = 1$, so that the mass m_i represents the interaction with i of a centre located at the unit distance from i , per attractiveness unit. Sheppard, on the other hand, expresses a_c as a function of the attractiveness of c , but we are not concerned by the definition of attractiveness nor by the expression of such a function, so that we define a_c itself as the attractiveness.

Thus presented, the GG model seems able to predict the total volume of interaction originating in a customer location (T_i) or emanating from a centre (P_c). It has been understood in this way by Thill (1985), for instance, who drew isolines of total interaction potential in Belgium (see our Section 9.2), and by Sheppard (*op. cit.*). However, in many models fitting the GG model as far as interaction is concerned, masses (or the factor K ; less often attractivenesses) are endogenous; see our Sections 1.4, 1.5, and 1.6. Many empirical or theoretical studies trying to validate the GG model have in fact focused on the way that model distributes interaction, not on the volume of interaction. This is not surprising, since the pioneering works on gravity models, like that of Reilly (1929, 1931), only attempted to explain the ratio between the interactions with two centres.

The function $F(\delta)$ is frequently set equal to the Pareto function, $\delta^{-\beta}$ (with $\beta > 0$), by geographers, like Illeris (1967), Berry and Lamb (1974), or Huff and Lutz (1979); or, more precisely, to $c^{-\beta}(\delta)$, transportation costs being denoted here as $c(\delta)$. The model is then the well-known *gravity model*, also called *logarithmic* or *Pareto model*. Tuominen (1949) and Godlund (1956) have shown that $Z_j \cap Z_k$ is a circle when $F(\delta) = \delta^{-\beta}$. The gravity model has received

support from the theoretical study of Niedercorn and Bechdolt (1969; see Section 1.4).

In opposition with this, theoretical (but not necessarily convincing) arguments as those of Schneider (1959) and Wilson (1970; see Section 1.4) indicate that $F(\delta) = \exp[-\beta c(\delta)]$. The same conclusion is drawn if we investigate the issue in the perspective of the multinomial logit model, as done by Beckmann (1976a; see Section 1.5). One may thus expect that a good candidate for $F(\delta)$ would be the negative exponential function, $\exp(-\beta\delta)$. But this, as well as the 'leptokurtic' ² (or square-root) exponential $\exp(-\beta\sqrt{\delta})$ and the normal or Gaussian function $\exp(-\beta\delta^2)$, is quoted by Berry and Lamb (*op. cit.*) as yielding weaker empirical adjustments than the Pareto function.

Beckmann (*op. cit.*) noted that the reconciliation of the two points of view consists in setting the equality $c(\delta) = \ln(\delta)$: in many cases indeed, mental perception is logarithmic. Nevertheless, the Pareto function itself is not really satisfying. It generally overestimates the influence of short distances and underestimates that of long ones; see Taylor (1971) ³. As underlined by Berry and Lamb, there is in fact no fundamental reason, *a priori*, for assigning $F(\delta)$ to any particular function. Hence the interest of the present work.

Coming back to the generalized gravity model, let us consider a set $J = \{j, k\}$ of two centres only. Here again, the set of places where the influence of k is not weaker than that of j is generally called the *market area* of k ⁴; we denote it also as Z_k . One might think of expressing this by the inequality $t_{ik} \geq t_{ij}$. However, all the points where $m_i = 0$ should then belong to $Z_j \cap Z_k$. To avoid this difficulty, it is better to resort to the notion of *potential* also introduced by Stewart (*op. cit.*); i.e., potential interaction by mass unit. When mass stands for population, for instance, the *partial potential* V_{ic} is the individual interaction with a centre c . The interaction with c may thus be written as

² A probability density is leptokurtic when it is more slender than the Gaussian. The term should thus be used for any distance deterrence function $\exp(-\beta\delta^a)$ if $a \in]0, 2[$, $\beta > 0$.

³ To try to improve the gravity model, the *log-normal* model uses the distance deterrence function $\delta^{-\beta \ln(\delta)}$, or more precisely $\exp[-\beta \max^2\{0, \ln(\delta)\}]$, with $\beta > 0$; see Taylor (*op. cit.*), Goux (1962).

⁴ Other names are in use, concerning the present section and the next ones in this chapter: influence area ('zone d'influence', 'zone de rayonnement'), tributary area, sphere of influence, (retail) trade area, trading area, predominant polarization area ('aire de polarisation prédominante'), etc.

$$t_{ic} = m_i V_{ic}.$$

Z_k is then defined, independently of the masses, as the set of points where

$$V_{ik} \geq V_{ij}.$$

In the generalized gravity model, we have

$$V_{ic} = a_c F(\delta_{ic}).$$

We thus face a problem that is, formally, totally similar to the one studied above in Section 1.1, with V instead of u . If we set, here also,

$$h = - \ln \circ F$$

and

$$p_c = - \ln(a_c) \quad \forall c \in C,$$

Z_k may be defined by inequality (1.1)⁵. In this last formulation, Z_k appears as the set of places i where the proximity of k wrt. that of j [expressed by the difference $h(\delta_{ij}) - h(\delta_{ik})$] surpasses or equals the advantage of j in terms of attractiveness (i.e., the index Q of comparative attractiveness). As to the reason of that formal identity with the all-or-nothing model, the reader is referred to our Sections 1.4 and 1.6.

Introducing the function h in the generalized gravity model and similar ones (i.e., adopting an additive formulation) simplifies the statements, as will be seen throughout this work. That function h is called hereafter the *transportation cost function*, because of the analogy with the first problem described in Section 1.1, and because it deserves that name in models like those of Schneider or Wilson, as well as in the multinomial logit model. But the reader should bear in mind that this is only one of the possible interpretations. In the model of Niedercorn and Bechdolt (*op. cit.*), $h(\delta) = \ln[c(\delta)]$.

⁵ This is not true, of course, when the distance deterrence function itself depends on a_c , and must thus be rewritten as $F(a_c, \delta_{ic})$; see however the end of this Section 1.1. When F depends on the mass m_i , it is even impossible to introduce the notion of potential. See Isard (1960, p. 512) for those extensions of the gravity model.

Notice also that the sign of $h(\delta)$ does not matter from our viewpoint. Inequality (1.1) shows that we may add any constant number to the function h without modifying the market areas. The positivity of $h(\delta)$ is thus meaningless; $h(\delta)$ appears only as a metric scale of transportation cost disutility. What is really important is the difference between $h(\delta)$ and the value of h at some reference distance; e.g., $h(1)$. Those remarks are particularly relevant when $h(\delta)$ is equal to $\ln(\delta)$ or $\ln[c(\delta)]$.

In fact, as the distance deterrence function in the initial formulation of the classical gravity hypothesis by Stewart (*op. cit.*) was $1/\delta$ or $1/c(\delta)$, it is the function $1/F(\delta)$ which is spontaneously considered to express a *perceived* or *behavioural distance* or transportation cost. It is also called the *resistance of distance* [see Smith (1975), pp. 144-145, and Beckmann (1971), p. 14].

1.3. Constrained generalized gravity models

In all the three 'constrained' versions of the GG model, the interaction t_{ic} remains proportional to $F(\delta_{ic})$. In the *consumption-constrained* (or *origin-constrained*) model, interactions are given by

$$t_{ic} = \mu_i V_{ic} ,$$

where the coefficients μ_i (playing the role of unknown or variable masses) are determined by the constraints

$$\sum_{c \in C} t_{ic} = T_i \quad \forall i \in M ,$$

where T_i is the total interaction, known a priori or depending only on the attributes of the consumer location i , between the place i and the two centres; M is the set of customers' locations. Of course, there is no difference between that model and the genuine GG model when we compute the market areas, since the coefficients μ_i have no influence on them. Notice, however, that this constrained model does not predict the total consumption amount T_i but only the way it is distributed among the centres, and that it is thus formally identical with the multinomial logit model model; see below, Section 1.6.

In the *production-constrained* (or *destination-constrained*) GG model, the potentials, not the masses, are affected. Potential is now written as

$$V_{ic} = \alpha_c F(\delta_{ic})$$

and the constraints determining α_j and α_k (appearing as unknown or variable attractiveness coefficients) are

$$\sum_{i \in M} t_{ic} = P_c ,$$

where P_c is the total production, known a priori or related to the attributes of c only, of the centre c . Our study applies to this model; however, changes in the deterrence function F , in the masses m_i , in the set of consumer locations M , or in the total productions P_c , imply changes in α_j and α_k and thus in Q and Z_k , since

$$\alpha_c = P_c / \sum_{i \in M} m_i F(\delta_{ic}) .$$

Notice that, consequently, Z_k is not here a market area as meant so far: it does not tell us whether a point i of $Z_k - M$ would be more influenced by k or by j if its mass m_i were nonzero. The only set that really matters is $Z_k \cap M$, the real market area.

The commentary on the production-constrained model remains true when the model is of the third kind: the *doubly constrained* GG model; which means that it is at the same time consumption- and production-constrained, with

$$t_{ic} = \mu_i \alpha_c F(\delta_{ic})$$

and the obvious compatibility condition

$$\sum_{i \in M} T_i = \sum_{c \in C} P_c .$$

The last equality, however, makes it very difficult to study the influence of a change in the sets C and M .

1.4. Models attempting to justify the generalized gravity model

Another category is constituted by attempts to justify the GG model on a theoretical level; the reader is referred to Sen and Smith (1995) for an exhaustive review. It seems that all those models are more or less different from the genuine GG model. Among them, the closest to the GG model is probably Niedercorn and Bechdolt's (*op. cit.*). The attractiveness, a_k , is here replaced by a coefficient of the form $a_k / \sum (a_c : c \in C)$, which implies that it is modified by any change in the set C . A similar characteristic appears in the model of Choukroun (1975), where it also affects the masses m_i , replaced by $m_i / \sum (m_{i'} : i' \in M)$. However, the market areas are the same if we maintain the attractivities a_k and the masses m_i . Niedercorn and Bechdolt's model, at least, can thus be directly studied in our perspective.

Wilson (1967, 1970) and Choukroun, by respectively maximizing entropy⁶ and utility, obtained a model that looks very much like the GG one. The function corresponding to the distance deterrence function, $F(\delta_{ic})$, is here $\exp(-\mu c_{ic})$, where c_{ic} is the transportation cost between i and c ; to apply our study, we just have to replace c_{ij} by $c(\delta_{ic})$, where c is the transportation cost function *stricto sensu*. However, the coefficient μ is not given *a priori* as in the GG model. It is, on the contrary, the Lagrange multiplier associated with a constraint on total transportation costs which is absent from the GG model. (Incidentally, this shows that the striking resemblance with the multinomial logit model, studied below in Section 1.6, is merely casual). The Wilson-Choukroun model may be submitted to our study, and we know *a priori* to which family of curves belongs the indifference line between two centres, j and k , as we may take c as (formal) transportation cost function: its equation has the form

$$c(\delta_{ij}) - c(\delta_{ik}) = Q / \mu ,$$

for some constant Q . But, as we have said, μ is endogenous: we have to compute it, and it depends on all the data of the problem. (It should be also noticed,

⁶ Instead of maximizing consumer utility as the other authors mentioned in this section, Wilson considers the situation *a posteriori*. The interaction t_{ij} representing here the number of trips from i to j , he assumes this minimal information: on the one hand, the total transportation costs; on the other, the total number of trips (P_j) ending in each centre (j), and the transportation cost corresponding to each origin-destination pair. Under those constraints, he then maximizes entropy. This amounts to looking for the most probable distribution of trips; the probability of $t = (t_{ij} : i \in M, j \in J)$ is supposed to be proportional to the number of possible arrangements of trips yielding t .

though it bears no influence on the market areas, that the masses are not necessarily present in Wilson's model, but may be introduced as endogenous variables). The model of Golob et al. (1973) displays the same difficulty, besides the fact that μ is here the common value of a series of Lagrange multipliers that must happen to have the same level. However, it will be seen later that the knowledge of function h in (1.1) often suffices for predicting the main properties of the market areas. Our work is thus easily applied to those problems.

The second model proposed by Choukroun (*op. cit.*) offers a still bigger difficulty. It introduces a probability distribution of the total transportation costs (and of the total interaction, on which bears a constraint that we have not mentioned), so that $F(\delta)$ is the average of $\exp[-\mu c(\delta_{ic})]$, μ itself being now stochastic. For us, the situation is thus much worse than before: the equivalent of function h in (1.1) is not known *a priori*, and subject to change whenever any element of the problem is modified.

Last, notice that a significant difference between the GG model and the models of Choukroun, Wilson, and Golob et al., as well as between the GG model and two of its constrained versions, is that the market areas depend on M , the consumer location set.

1.5. Other models reducible to GG or all-or-nothing models

A last category consists of interaction models that are, only formally, reducible to GG models. Such is a linear interaction model suggested by Isard (1960, p. 512), from which we may derive a general additive interaction model where potential takes the form

$$V_{ic} = - p_c - h(\delta_{ic}) ,$$

in which p is related with the attributes of centre c , and h is some increasing function of distance, with the condition that

$$p_c + h(\delta_{ic}) \leq 0 ,$$

on the range of distance considered ⁷. Once submitted to an exponential transformation, Isard's formula is formally identical to that of the GG model as far as the main problem is concerned. As to the constrained models, the conclusions are not altered, though the formulæ are.

Allen and Sanglier (e.g. 1981) replace the deterrence function F of the gravity model by a centre-dependent Pareto-like function F_c , so that potential becomes

$$V_{ic} = a_c F_c (\delta_{ic})$$

with $F_c (\delta_{ic}) = (p_c + r_c \delta_{ic})^{-\beta}$,

for some $\beta > 0$; p_c is the mill price at c ; r_c is the transportation rate related to c . The price, p_c , is thus not integrated into the attractiveness coefficient a_c . Let us generalize that model in the following way:

$$F_c (\delta_{ic}) = [p_c + r_c h (\delta_{ic})]^{-\beta},$$

with the condition that

$$p_c + r_c h (\delta_{ic}) > 0$$

on the considered range of distance; $r_c h(\delta)$ may represent the transportation cost. In the two-centre problem, the market area of k becomes defined by

$$[p_j + r_j h (\delta_{ij})] / [p_k + r_k h (\delta_{ik})] \geq (a_j / a_k)^{1/\beta},$$

to which we may give the form

$$(r_j/a_j^{1/\beta}) h (\delta_{ij}) - (r_k/a_k^{1/\beta}) h (\delta_{ik}) \geq p_k/a_k^{1/\beta} - p_j/a_j^{1/\beta},$$

showing that we face an inequality similar to (1.2).

Last, let us mention Gillen and Guccione (1993) who start from an all-or-

⁷ The interaction formula suggested by Isard and to which we allude is $t_{ic} = k - \beta \delta_{ic}$, where k and β are positive and constant. He also proposes (*loc. cit.*) two other formulations that we can write as: $t_{ic} = m_i a_c - \beta \delta_{ic}$, and $t_{ic} = k \ln(m_i a_c)/\delta_{ic}^\beta$. It is only when m_i is the same for all customers $i \in M$ that we can apply our study to those formulæ.

nothing model of consumer behaviour into which they introduce incomplete price information and search costs. Space then appears structured into several regions described by inequalities similar to (1.1).

1.6. The multinomial logit model

In a very general context, the customer i confronted with several alternatives will make his choice by comparing their utilities. Let u_{ij} be the utility for the customer i to make the choice j , and suppose that u_{ij} is the sum of a constant term $\hat{u}_{ij}$ and a stochastic term X_i :

$$u_{ij} = \hat{u}_{ij} + X_i .$$

This term may express a stochastic behaviour of the individual consumer, as well as a lack of precision in our data, an extreme case being that of nonmeasured data; if the index i simply refers to the location of the customer, X_i may express differences in decision rules between the consumers located there.

If the X_i 's are independent Gaussian variables with zero average and a same standard deviation σ , the model is the *multinomial probit model*. It is well-known that if the normal variables are approximated by Gumbel ones having the same average 0 and standard deviation σ , the probability p_{ij} that i chooses j rather than any other alternative $c \in J$ (the set of all alternatives) is given by

$$p_{ij} = \exp (\mu \hat{u}_{ij}) / \sum_{c \in C} \exp (\mu \hat{u}_{ic}), \quad (1.4)$$

where $\mu (= \pi/\sigma\sqrt{6})$ is the scale parameter or distribution dispersion parameter; see e.g. Ben-Akiva and Lerman (1985), pp. 103-106. The model so obtained is called the *multinomial logit model*. The Gumbel (or Gnedenko, or Weibull, or 'double negative exponential', or 'extreme value') distribution is close to the normal (Gauss) distribution. Its cumulated density function is characterized by μ , and by the mode, m :

$$\Pr [X \leq x] = \exp \{-\exp [-\mu(x - m)]\} .$$

The average is $m + \gamma/\mu$, γ being Euler's constant, approximately equal to 0.5772 ⁸.

Coming back to our subject and, more precisely, to the all-or-nothing model, we set

$$\hat{u}_{ij} = -h(\delta_{ij}) - p_j .$$

The set of alternatives being now the set of centres C , we see that, concerning (1.4), we have

$$\exp(\mu \hat{u}_{ij}) = \exp(-\mu p_j) \exp[-\mu h(\delta_{ij})] ;$$

or, if we set $a_j = \exp(-\mu p_j)$ and $F(\delta_{ij}) = \exp[-\mu h(\delta_{ij})]$,

$$\exp(\mu \hat{u}_{ij}) = a_j F(\delta_{ij}) .$$

The expression of p_{ij} [which is also the market share of j , as first emphasized in this field by Huff (1963) and, more generally, by Carey (1858-9)] is thus similar in the generalized gravity (GG) and multinomial logit (MNL) models. This has been brought to light by Beckmann (1976a) and Sheppard (*op. cit.*), although they did not mention the multinomial logit model ⁹. This model thus also gives rise to market areas described by (1.1).

Of course, we have here

$$t_{ij} = T_i p_{ij} ,$$

in which T_i (which cannot be predicted by the model) is the total interaction originating in location i . If we further identify that expression of t_{ij} with the GG model, we have to set

$$m_i = T_i / \sum_{c \in C} a_c F(\delta_{ic}) .$$

⁸ Euler's constant, usually denoted by γ , is the limit of the expression $\sum(1/i: i \in \mathbb{N} \cap [1, n]) - \ln(n)$ when n tends toward infinity.

⁹ Beckmann's X 's remain Gaussian, but he approximates the expression of p_{ij} by a logistic function. This would be, as a matter of fact, the exact expression of p_{ij} if the X_{ij} 's were Gumbel variables.

Masses are thus here generally endogenous variables, unless T_i also is endogenous and proportional to $\sum_{C \in C} a_c F(\delta_{ic})$. If T_i is exogenous, the model is identical with the consumption-constrained GG model; see Section 1.3.

1.7. Derived issues

Come back to our two centres: $C = \{j, k\}$. The area where the sales ratio or potential ratio V_{ik}/V_{ij} (equal, as said here above, to t_{ik}/t_{ij} if $m_i \neq 0$), is at least equal to some value σ is defined by

$$h(\delta_{ij}) - h(\delta_{ik}) \geq Q + \ln(\sigma),$$

and thus equal to $Z_k[Q + \ln(\sigma)]$; see Beckmann (1971, 1976a).

Other areas are described by the same formula. The area where the market share $V_{ik}/(V_{ik} + V_{ij})$ of k in i , equal as said above to the probability p_{ik} , is at least equal to τ (< 1) is described by the above-mentioned equality if $\sigma = \tau/(1 - \tau)$. We call that region the τ -share area of k ; see Sections 6.2 and 7.4, and Chapter 10. In this interpretation the demarcation set described by $h(\delta_{ij}) - h(\delta_{ik}) = Q + \ln(\sigma)$ is called an *isoshare line*. As to the area where the difference in market share or relative difference in potential $(V_{ik} - V_{ij})/(V_{ik} + V_{ij})$, is superior or equal to some fixed number $\zeta \in]-1, 1[$, we have $\sigma = (1 + \zeta)/(1 - \zeta)$.

Another issue is uncertainty. If, in the context of the all-or-nothing model, we consider the area where

$$u_{ik} + \varepsilon \geq u_{ij} \tag{1.5}$$

for some $\varepsilon > 0$, we see that this inequality is equivalent to

$$h(\delta_{ij}) - h(\delta_{ik}) \geq Q - \varepsilon,$$

so that the area is equal to $Z_k(Q - \varepsilon)$.

That area may be viewed as the extended market area of k . The margin ε expresses our uncertainty about utility values, due to behavioural variations

between consumers or to imprecision in our knowledge. This is the case, for instance, in a problem similar to that considered in Section 1.6, if X_i is uniform over $[-\varepsilon, \varepsilon]$. The consumers are supposed to purchase in k without any doubt if they are located inside $Z_k(Q + \varepsilon)$, which we may call a *strict market area* of k ; whereas their behaviour is uncertain inside $Z_k(Q - \varepsilon) - Z_k(Q + \varepsilon)$, the *uncertainty area* or *shared influence area*, corresponding to the inequality $|h(\delta_{ij}) - h(\delta_{ik}) - Q| \leq \varepsilon$; see Chapter 10.

Needless to say, if we replace (1.5) by the inequality $\acute{u}_{ik} \geq \eta \acute{u}_{ij}$, or for the GG and MNL models by $t_{ik} \geq \eta t_{ij}$, we get nearly the same results through the use of a logarithmic transformation.

We may also define an extended market area by the inequality

$$u_{ik} \geq \eta u_{ij}, \quad \text{for some } \eta \in]0, 1[. \quad (1.6)$$

The associated shared influence area is defined by $u_{ij}/\eta \geq u_{ik} \geq \eta u_{ij}$. Inequality (1.6) is equivalent to

$$\eta h(\delta_{ij}) - h(\delta_{ik}) \geq p_k - \eta p_j,$$

which is similar to (1.2).

Last, defining an extended market area by

$$\acute{u}_{ik} + \varepsilon \geq \acute{u}_{ij},$$

with $\varepsilon > 0$ (or by the same inequality with $\acute{u}$ replaced by t in the GG and MNL models), yields the following inequality, according to (1.3):

$$a_k F(\delta_{ik}) - a_j F(\delta_{ij}) \geq -\varepsilon,$$

which is again formally similar to (1.2), h being replaced by $(-F)$. Remember indeed that the function g in (1.2), like h in (1.1), is not required to be positive.

*

A last remark. In all the models we have presented in this chapter, the

writing suggests that the set M of consumer places where $m_i \neq 0$ is finite, or at least discrete; but it is not necessarily so. In particular, this set may be two-dimensional. In that case, m_i should better be called a *density*, and t_{ij} , a *interaction density*; and the total production of a centre would be yielded by an integral instead of a sum. But, beyond that purely formal question, another remains. When we face, in reality, a finite and known set M of consumer locations, or when the model itself makes use of such a finite set (as may be and generally is the case in the production- and doubly constrained models studied hereafter, for instance), what is the meaning of the properties studied in this work ? We shall come back to that issue in Section 5.6 and in Chapter 10.

2. A few introductory properties

2.1. Delimitation of the two-centre problem and definitions

Given some increasing real function h ('transportation cost function', in short: t.c.f.), a set of two centres: $C = \{j, k\}$, and two real numbers associated with those centres: p_j and p_k ('mill prices'), our problem is to describe the market areas of j and k . If we start with a multiplicative all-or-nothing model or with a distributive one (GG, MNL), we set $p_c = -\ln(a_c)$ and $h = -\ln \circ F$. As a matter of fact, as two centres only are considered, the difference between mill prices only, Q ('attractiveness index'), plays a part, and the market areas may be written as follows:

$$\begin{aligned} Z_j &= \{i: h(\delta_{ij}) - h(\delta_{ik}) \leq Q\} \\ Z_k &= \{i: h(\delta_{ij}) - h(\delta_{ik}) \geq Q\}, \end{aligned} \tag{2.1}$$

where δ_{ij} and δ_{ik} are the Euclidean distances between i and j and between i and k , respectively.

The issue, a purely geometrical one, is clearly akin to the once fashionable studies about curves as recollected in the treatises of Salmon (1879) or Gomes Texeira (1908). However, our approach is rather different. In particular, those authors do not seem to have paid much attention to questions considered in our Chapters 4 and 5 (see e.g. the rash discussion about inflexion points in Salmon, item 204).

From now on we suppose 1° that function h is continuous on $]0, \infty[$, right-continuous *lato sensu* at 0, i.e.,

$$h(0) = \lim_{\delta \rightarrow 0, \delta > 0} h(\delta);$$

and 2° strictly increasing. The former property ensures that Z_k and Z_j are closed sets; the latter, that $Z_j \cap Z_k$ is one-dimensional when $\neq \emptyset$. Hurter and Lowe (1976, p. 176) already pointed out that the absence of 'plateau' segments in function h implies this one-dimensionality (those issues are developed in Hanjoul et al., Section 12).

.2.2.

Without loss of generality, we assume that the attractivity of centre j is greater than that of centre k ; i.e.,

$$Q > 0 .$$

As the area Z_k is then obviously smaller than Z_j , it seems most of the time convenient to focus on the properties of Z_k without mentioning the complementary ones of Z_j . On the other hand, some properties are easier to enounce in reference to $Z_j \cap Z_k$, the indifference zone, for which we have

$$Z_j \cap Z_k = \{i: h(\delta_{ij}) - h(\delta_{ik}) = Q\}.$$

Notice that $Z_j \cap Z_k$ is often the common boundary of Z_j and Z_k , but not always; see Fig. 5.3, where $Z_j \cap Z_k$ is equal to the boundary of Z_k , $b(Z_k)$, but strictly contains $b(Z_j)$. Giving a wider meaning to a technical term of Hyson and Hyson (1950) and Hurter and Lowe (1976), who used it for Cartesian ovals only, we call that set a *hypercircle*; and j and k , its *foci*.

To make the writing more concise, some compact notation is used. The index i is dropped when this entails no confusion, so we write δ_j , δ_k instead of δ_{ij} , δ_{ik} ; $h(\delta_{ij})$ and $h(\delta_{ik})$ are simplified into h_j and h_k ; and similarly for the derivatives of h . Also, we set

$$\Delta h_i = h(\delta_{ij}) - h(\delta_{ik}) ,$$

usually shortened into Δh and called the transportation cost difference.

In order to be able to describe accurately the position of any point w.r.t. the centres, we assimilate the plane to $\mathbb{R}^2$; $\mathbb{R}$ being the set of all real numbers, $]-\infty, \infty[$. The point $(0,0)$, called o , is the middle of the segment $[jk]$. The centres themselves are $j = (-\delta_{ok}, 0)$ and $k = (\delta_{ok}, 0)$, with the exception of Proposition 2.6.a. The set of points not more distant from j (resp. from k) than from k (resp. from j) is thus $\mathbb{R}_- \times \mathbb{R}$ (resp. $\mathbb{R}_+ \times \mathbb{R}$). The straight lines are referred to as *vertical* lines when parallel to the y -axis, *horizontal* lines when parallel to the x -axis, and *oblique* lines otherwise. We also attach to any point p of the plane a system of polar coordinates (δ_p, φ_p) , the angle φ_p (abbreviation for φ_{pi}) being defined as in Fig. 2.1.

Last, to avoid any confusion with a negative power, we denote the inverse relation of any function or relation f by $f^{\wedge}$.

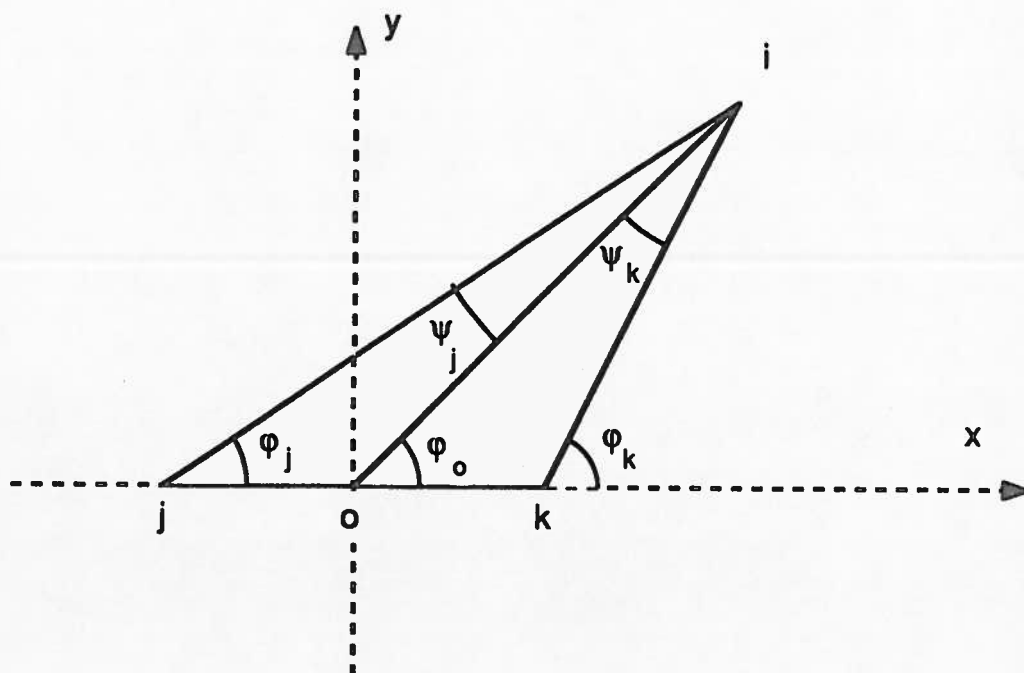


Fig. 2.1. Conventions for the 2-centre case study.

2.2. The identification problem

A first question, which is not related to the particular space and distance considered, is whether it is possible to deduce the transportation cost function from the knowledge of the market areas. Obviously, that identification can usually not be complete, as the set of distances concerned by the equation of $Z_j \cap Z_k$ (from which that identification should be carried on) is generally not $[0, \infty[$. The determination of that set is made in Corollary 4.2.

On the other hand it is obvious from (2.1) that replacing $h(\delta)$ by $\mu h(\delta) + \lambda$ has no effect on market areas if $\mu > 0$ and Q is replaced by μQ . In particular, constant fixed transportation costs bear no influence on market areas. (Strictly speaking, however, their presence makes it more probable that each centre belongs to its own market area if we maintain $h(0) = 0$, thus dropping the continuity property of h ; see Hanjoul et al., Section 12.2).

But function h may be transformed more deeply without the market areas being altered. Let us denote indeed by Z_j and Z_k (respectively, by $Z^{\sim}_j$ and $Z^{\sim}_k$) the market areas associated with a transportation cost function h (resp., $h^{\sim}$) and a value Q of the attractivity constant. Let, for any $c \in \{j, k\}$,

$$I_c = \{\delta_c: i \in Z_j \cap Z_k\} ;$$

and set $\mathbb{N}^* = \mathbb{N} - \{0\}$, with $\mathbb{N}$ equal to the set of all natural numbers; for any function Φ and any set E , we define $\Phi\langle E \rangle$ as $\{\Phi(\xi): \xi \in E\}$. And the function Φ is periodic on E iff., for some $\tau > 0$ and for every $\xi \in E$, $\Phi(\xi + \tau) = \Phi(\xi)$; the minimal value of τ is the period, T , of Φ on E . The general property is then:

(2.1) (a) If h and $h^{\sim}$ are related as follows:

$$h^{\sim}(\delta) = h(\delta) + f[h(\delta)] , \quad \forall \delta \geq 0, \quad (2.2)$$

where f is some periodic function with period $T > 0$, and if $Q = nT$ for some $n \in \mathbb{N}^*$,
then $Z_j \cap Z_k = Z^{\sim}_j \cap Z^{\sim}_k$;

(b) If $Z_k = Z^{\sim}_k$ and $Z_j = Z^{\sim}_j$,

then h and $h^{\sim}$ are related by (2.2) $\forall \delta \geq 0$, for some function f periodic on $h\langle I_k \rangle$, with period Q/n for some $n \in \mathbb{N}^*$.

Take indeed any $i \in Z_j \cap Z_k$. We thus have $h_j - h_k = Q = nT$, so that $f(h_j) = f(h_k)$. Hence,

$$h^{\sim}_j - h^{\sim}_k = h_j - h_k + f(h_j) - f(h_k) = Q + 0 = Q ,$$

which means that $i \in Z^{\sim}_j \cap Z^{\sim}_k$; and thus that $Z_j \cap Z_k \subseteq Z^{\sim}_j \cap Z^{\sim}_k$.

Take now some $i \in Z^{\sim}_j \cap Z^{\sim}_k$, and consider the number $\xi^* = h_k + Q$. According to the definition of $h^{\sim}$ and f , we thus have $f(\xi^*) = f(h_k)$ and thus

$$\xi^* + f(\xi^*) = h_k + Q + f(h_k) = h^{\sim}_k + Q = h^{\sim}_j ;$$

hence, $\xi^* + f(\xi^*) = h_j + f(h_j)$. It is clear that the function g , defined by the equality $g(\xi) = \xi + f(\xi)$, must be strictly increasing w.r.t. ξ on $h\langle\mathfrak{X}\rangle$, since $h^\sim$, equal to $g \circ h$, is strictly increasing on $[0, +\infty[$ as being a t.c.f. The equality $g(\xi^*) = g(h_j)$, in which ξ^* and h_j belong to $h\langle\mathfrak{X}\rangle$, thus implies that $h_j = \xi^*$; i.e., $h_j = h_k + Q$, which means that $i \in Z_j \cap Z_k$. This shows that $Z_j^\sim \cap Z_k^\sim \subseteq Z_j \cap Z_k$; which result, combined with the reversed inclusion established above, proves the equality of the two hypercircles.

Now for item (b). As h is strictly increasing, its inverse relation $h^\sim$ is a function; we may thus define a function f on $\mathfrak{X}$ by $f(\xi) = h^\sim[h^\sim(\xi)] - \xi$; so that, for any δ , $f[h(\delta)] = h^\sim(\delta) - h(\delta)$. Take now some $\xi \in h\langle I_k \rangle$. By definition, there is some i belonging to both $Z_j \cap Z_k$ and $Z_j^\sim \cap Z_k^\sim$, and for which $\xi = h_k$. This yields $f(\xi) = h_k^\sim - h_k$, and, as $i \in Z_j \cap Z_k$,

$$\begin{aligned} f(\xi + Q) &= h^\sim[h^\sim(h_k + Q)] - (h_k + Q) \\ &= h^\sim[h^\sim(h_j)] - h_j \\ &= h_j^\sim - h_j. \end{aligned}$$

This entails that

$$f(\xi + Q) - f(\xi) = h_j^\sim - h_j - h_k^\sim + h_k = Q - Q = 0.$$

The function f has thus the announced periodicity property. **Q.E.D.**

Prop. 2.1 shows that the identification of the transportation cost function (which identification should probably be carried out by means of Fourier or Lagrange transforms) does not produce a unique result unless some assumptions are made. What such assumptions should be is a problem we leave out of the field of our investigation, but we may point out the fact that the requirement that function h should be concave would not suffice; the importance of concave t.c.f.'s will be seen throughout the present work. It is easily verified that the function $h^\sim$ defined as $\ln(\delta) + \alpha \sin[\ln(\delta)]$, for instance, is concave and strictly increasing if $\alpha \leq \sqrt{2}/2$; it yields the same hypercircles as the function $\ln$ when $Q = 2\pi n$, for all $n \in \mathbb{N}$.

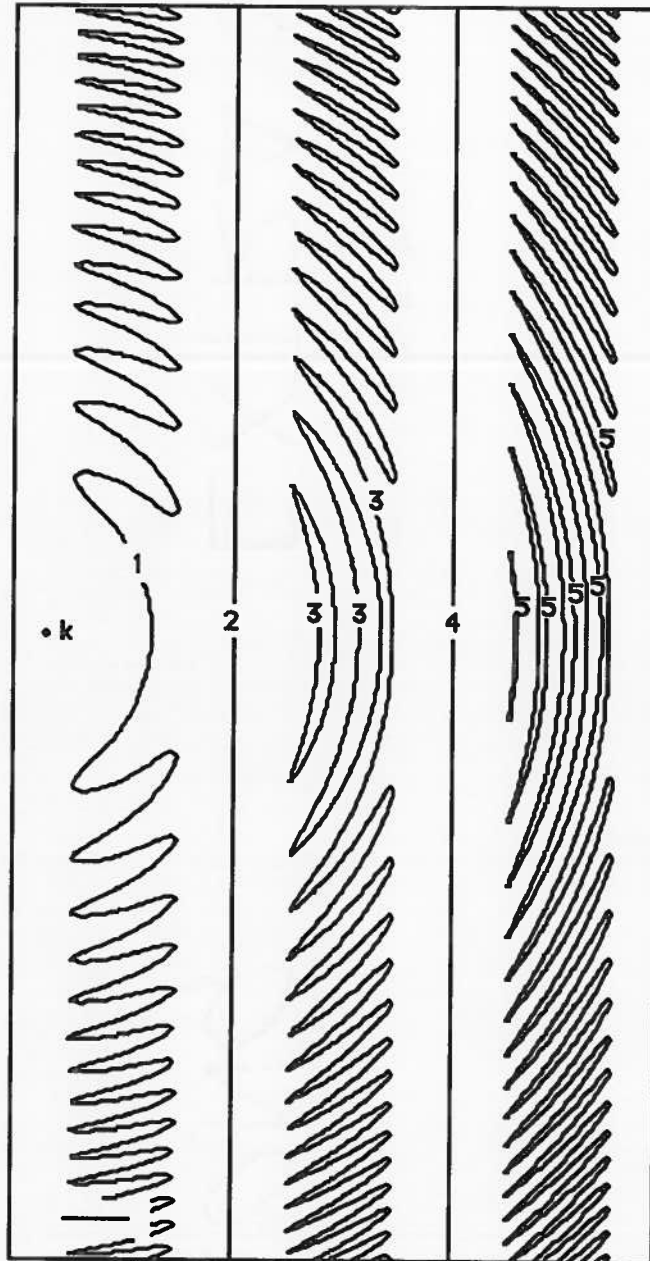


Fig. 2.2 — Market areas generated by the t.c.f. $\delta^2 + \sin(\delta^2)$ when $\delta_{jk} = 0.96$ and $Q = n\pi$, for indicated values of n ¹.

Let us illustrate Prop. 2.1 by some striking example. The function h defined by $h(\delta) = \delta^2 + \sin(\delta^2)$ is strictly increasing and we may choose it as t.c.f. It is known (see Prop. 2.7c) that when the transportation costs are given by δ^2 , the hypercircle $Z_j \cap Z_k$ is a vertical line. Consequently,

¹ In this figure as in most of the figures of Part I, centre j is symmetric of centre k w.r.t. the left edge of the frame; and the horizontal stroke down to the left is the distance unit.

this function h also yields such a vertical indifference set when $Q = 2\pi n$, for all $n \in \mathbb{N}$; see Fig. 2.2.

2.3. Position

The following proposition is obvious:

- (2.2) Market areas Z_j and Z_k are symmetric with respect to the x -axis $\mathcal{R} \times \{0\}$.

The properties of Z_k and Z_j above the x -axis consequently characterize the whole market areas. This point will be used throughout the paper. It is also worth noticing that the couple (δ_j, δ_k) constitutes a system of coordinates of $\mathcal{R} \times \mathcal{R}_+$ in this sense, that to every point of $\mathcal{R} \times \mathcal{R}_+$ corresponds exactly one couple (δ_j, δ_k) and that every such couple is associated with one point at most in $\mathcal{R} \times \mathcal{R}_+$. Following Lockwood (1961), we call *bipolar* those coordinates ². More generally, (h_j, h_k) constitutes a system of coordinates since h is strictly increasing and thus injective; we shall refer to them as *generalized bipolar*.

- (2.3) Market area Z_k is strictly included in the right-hand open half plane at the right of the y -axis; i.e., $Z_k \subset \mathcal{R}_+^* \times \mathcal{R}$.

This derives from (2.1), where the fact that $Q > 0$ entails that $\delta_j > \delta_k$ in Z_k . The property implies that $\mathcal{R}_- \times \mathcal{R} \subset Z_j$, which means that any point i is more influenced by j than by k when both $p_j < p_k$ (i.e., $Q > 0$) and $\delta_j \leq \delta_k$ (i.e., $i \in \mathcal{R}_- \times \mathcal{R}$).

- (2.4) Let i be a point moving on $Z_j \cap Z_k$. The distances from i to all the points of $[jk]$ increase or decrease simultaneously.

² Jones (1979) uses the word 'dipolar', which should best be reserved for another use; see Hanjoul et al., Section 11. Hanjoul and Thill (1987) call them *bicentric* coordinates.

The proof is as follows. First, δ_j and δ_k increase or decrease simultaneously since $\delta_j = h^*[h(\delta_k) + Q]$ on $Z_j \cap Z_k$; h and h^* are indeed strictly increasing functions. Second, take any point $p \in [jk]$. The following relations hold:

$$\delta_j^2 = \delta_{jp}^2 + \delta_p^2 + 2 \delta_{jp} \delta_p \cos(\varphi_p)$$

$$\delta_k^2 = \delta_{kp}^2 + \delta_p^2 - 2 \delta_{kp} \delta_p \cos(\varphi_p) .$$

Eliminating $\cos(\varphi_p)$ makes δ_p^2 appear as an increasing function of both δ_j^2 and δ_k^2 ; hence the property. Q.E.D.

2.4. Dependence on Q and δ_{jk}

The next two propositions deal with the influence of two basic data of the problem: the attractivity constant Q and the distance δ_{jk} between the two centres. They show that the market area of the less attractive centre, Z_k , shrinks when the difference in attraction between the centres increases or when the distance between them decreases.

(2.5) *Market area Z_k is a strongly decreasing function of Q .*

The meaning of '*strongly decreasing*' is as follows. Suppose that Z'_j and Z'_k are the market areas corresponding to another value Q' of the attractivity constant. The property (see Fig. 2.3) signifies that

$$Q' < Q \Rightarrow Z_k \subseteq Z'_k - Z'_j .$$

Otherwise said, if parameter Q decreases, any point of Z_k remains in Z'_k (as $\Delta h \geq Q \Rightarrow \Delta h > Q'$), and the points of the previous indifference zone $Z_j \cap Z_k$ cannot remain on the new one (as $\Delta h = Q \Rightarrow \Delta h > Q'$). This is immediately understandable from the fact that the hypercircles corresponding to various values of the attractivity constant Q are *iso- Δh lines* of the plane.

Proposition 2.5 may be particularized to the breaking points of the x -axis. Denote for instance as ℓ and r the extreme left and right breaking

points, if any; then: x_l is a strictly increasing, and x_r a strictly decreasing, function of Q .

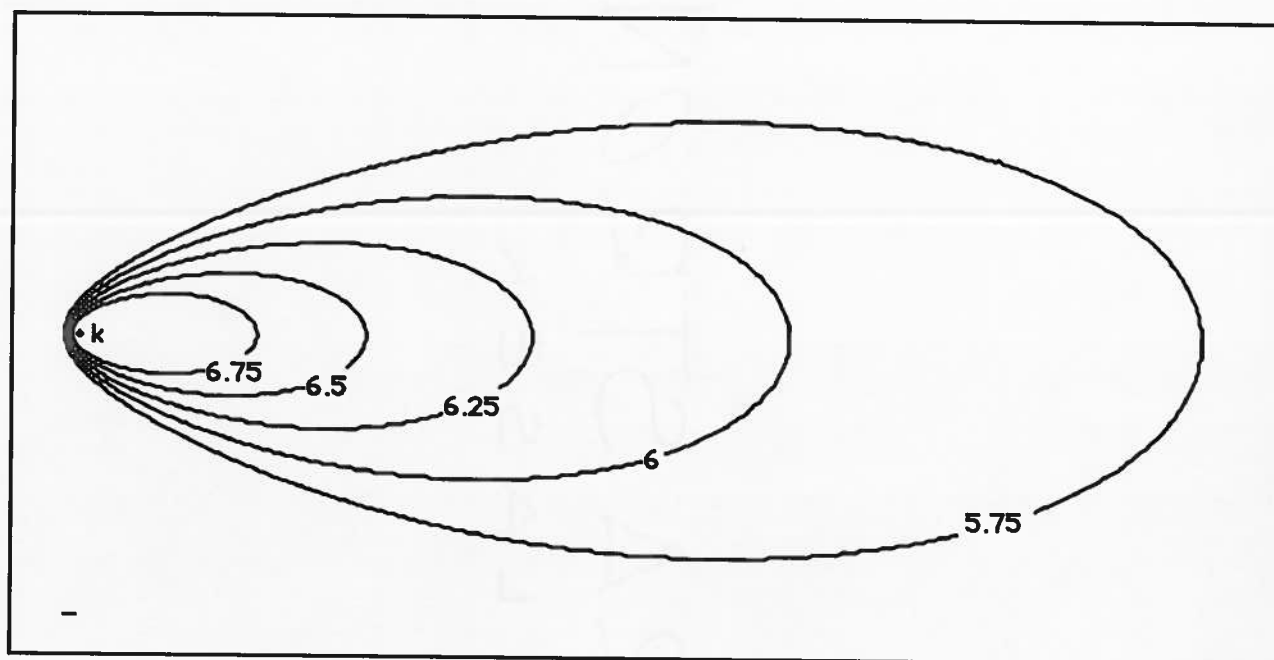


Fig. 2.3 — Market areas generated by the power t.c.f. $\delta^{0.9}$
when $\delta_{jk} = 10$, Q being as indicated.

- (2.6) (a) If k is fixed and j is allowed to move on the x -axis at the left of k , then Z_k is a strongly increasing function of δ_{jk} ;
- (b) Whatever the way the positions of the centres vary, the measure $|Z_k|$ of area Z_k is a strictly increasing function of δ_{jk} when $|Z_k|$ is neither 0 nor $+\infty$.

Subproposition (b) is straightforwardly derived from (a), as the change in Z_k for a variation of the location of centres j and k may always be studied through some appropriate variable system of coordinates. As to point (a): consider the market areas for a given value $2L$ of the distance between centres. We have thus $j = (-L, 0)$ and $k = (L, 0)$. Let now j move toward the left, without modifying the system of axes. Point k remains equal to $(L, 0)$, whereas j has become $(-L', 0)$ with $L' > L$. Consequently, for all points of the half-plane $\mathbb{R}_+ \times \mathbb{R}$, the distance to j and thus the transportation cost difference Δh have strictly increased. In particular, as

$Z_k \subset \mathfrak{R}_+ \times \mathfrak{R}$ (Prop. 2.3), we have $(\Delta h)' > \Delta h \geq Q$ on Z_k , so that $Z_k \subseteq Z'_k - Z'_j$.
Q.E.D.

Notice that Prop. 2.6a may be particularized in the same way as Prop. 2.5a: x_l is a strictly decreasing and x_r a strictly increasing function of δ_{jk} if, k remaining fixed, j is allowed to move at the left of k on the x -axis.

The present section has detailed some results that do not depend in their essence on the transportation space under consideration, though they are enriched by the structure of the Euclidean plane. In contrast with this, we now come to a section in which we try to get a first insight into peculiar properties generated by that structure.

2.5. Some easy cases

The next proposition lists cases where Z_k is known to have a simple shape with well-established properties; 'coth' stands for 'hyperbolic cotangent', and 'sh' for 'hyperbolic sine'. The deterrence function $F(\delta)$, if we refer to the gravity model, then takes one of the forms: $\delta^{-\beta}$ (Pareto function), $\exp(-\beta\delta)$ (negative exponential case), $\exp(-\beta\delta^2)$ (normal or Gaussian case), $(K + \delta)^{-\beta}$, $(K + \delta^2)^{-\beta}$. As said in Sections 1.1 and 1.2, the first case has already been solved in the context of the GG model; the two sequent ones, in the all-or-nothing economic theory of market areas. Remember that β may be replaced by 1 in our perspective; 1_E is the identity function on a set E .

- (2.7) (a) If $h = \ln$, Z_k is an Apollonian disk relative to points j and k , with centre $[\delta_{ok} \coth(Q), 0]$ and radius $\delta_{ok} / \text{sh}(Q)$;
- (b) If $h = \ln_+$, Z_k is convex and limited by the right branch of a hyperbola with foci j and k ;
- (c) If $h = (.)^2$, Z_k is the half-plane $x \geq Q/2\delta_{jk}$;
- (d) If $h = \ln [K + (.)]$ for some $K > 0$, Z_k is convex, bounded, more extended in the direction of x than in that of y , and the indifference line is a Cartesian oval;

- (e) If $h = \ln [K + (.)^2]$ for some $K > 0$, Z_k is a disk, with centre $[\delta_{ok} \coth (Q/2), 0]$ and radius $\delta_{ok} / \sqrt{[sh^2 (Q/2) - K]}$.

All those properties are easily proved. As to (a), the inequality of Z_k may be written $\delta_j/\delta_k \geq e^Q$, which Apollonius of Perga has shown to describe a disk the centre of which lies on the x-axis at the right of k. Point (b) is obvious since the indifference line is described by $\delta_j - \delta_k = Q$, which clearly represents the right branch of a hyperbola, as also discovered by Apollonius. As regards point (c) [which would have already been exploited by Laguerre (+1886)], δ_j and δ_k can be expressed as functions of x and y, as follows:

$$\begin{aligned}\delta_j^2 &= (x + \delta_{ok})^2 + y^2 \\ \delta_k^2 &= (x - \delta_{ok})^2 + y^2\end{aligned}\tag{2.3}$$

from which appears the well-known relation

$$\delta_j^2 - \delta_k^2 = 2 x \delta_{jk} .\tag{2.4}$$

As the inequality concerning Z_k is $\delta_j^2 - \delta_k^2 \geq Q$, the proof of this item is complete. As to point (d), illustrated in Fig. 2.4, the reader is referred to the Chapter 13 of Hanjoul et al. The hypercircle, here the boundary of Z_k , is a Cartesian oval.

The last item, (e), is a consequence of item (a). Consider indeed the points j' and k' located above the plane at a distance $\sqrt{K}$ from j and k , respectively. The inequality of Z_k may then be written:

$$\begin{aligned}\ln (\delta_{j'}^2) - \ln (\delta_{k'}^2) &\geq Q ; \\ \text{i.e.,} \\ \ln (\delta_{j'}) - \ln (\delta_{k'}) &\geq Q/2 .\end{aligned}$$

In the 3-dimensional space, the market area of k would consequently be a sphere whose centre and radius are determined by (a), with Q replaced by $Q/2$. Z_k is the intersection of that sphere with the plane; which yields the result.
Q.E.D.

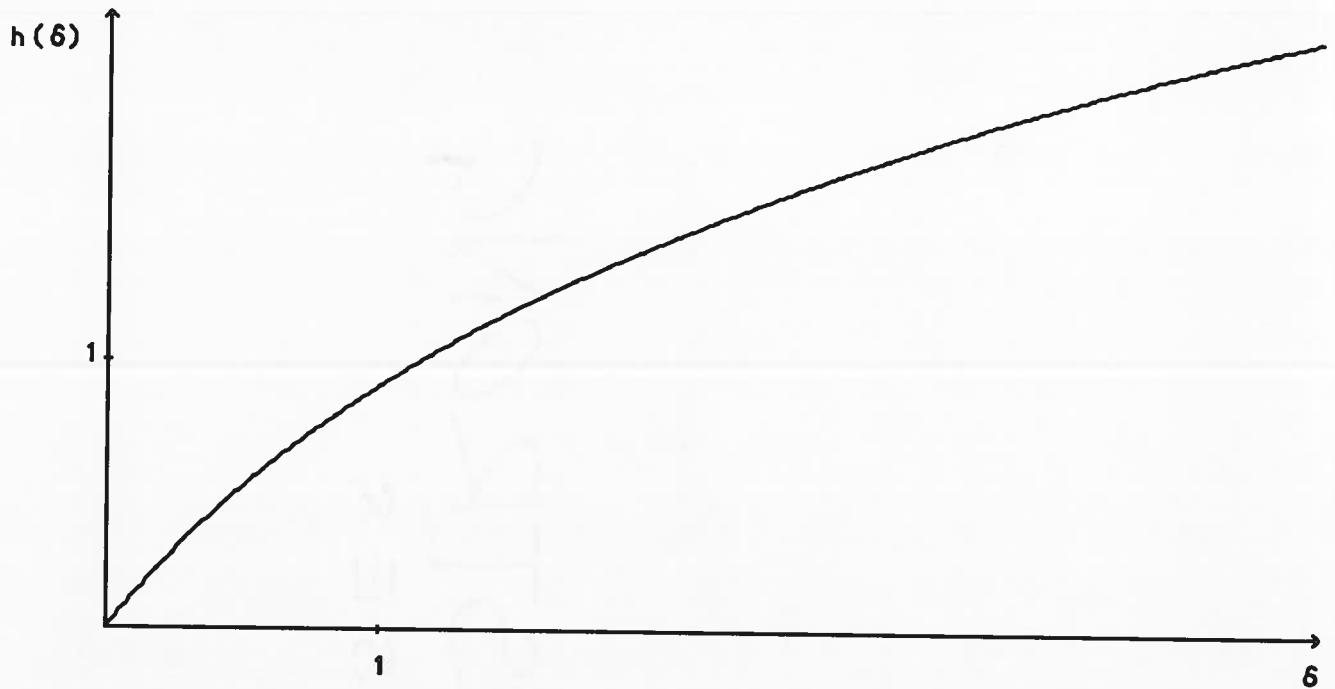


Fig. 2.4 — The t.c.f. $1.3 \ln(K+\delta)$, which generates Cartesian ovals as indifference lines ($K > 0$).

The value of K is 1.

Items (a), (d), and (e), call for two remarks. First: should we specify $K \geq 0$ instead of $K > 0$, item (a) would be a particular case of (d) and of (e). The centre (noteworthy independent of K) and radius mentioned in (e) are also valid for (a) if we substitute $Q/2$ by Q . Second: in (a) and (e), we know the radii of the disk, thus their superficies too. As far as we know, and apart from the approximate results of Hanjoul *et al.*, Chapter 11, this is the only case where such an analytic expression of the measure $|Z_k|$ is available.

This proposition provides us with a set of instances of the general properties we shall discover. Let us thus have a closer look at their main apparent features. In items (a), (b), and (d), function h is concave, which is empirically appealing if h is really intended to describe transportation costs. In item (b), h is linear; i.e., concave and convex at the same time. In those three cases, centre k lies in Z_k .

In item (c), function h is strictly convex. If we interpret $h(\delta)$ as the real transportation cost function $c(\delta)$, we are here in front of a situation

where a sum of short trips would be preferred to an equivalent long trip, at least if we neglect fixed costs. Usually, however, $c(\delta)$ is supposed to be concave, so that we may expect $h(\delta)$ to be concave; this is also true if $h(\delta) = \ln[c(\delta)]$. Nevertheless, Smith (1975, p. 146) views convex transportation cost functions as acceptable, and some empirical and theoretical arguments allow to doubt of their nonexistence [see Gatrell (1983), Huriot and Thisse (1986)]. Anyway, the interpretation of the GG model remains uncertain, as we have said in Section 1.2.

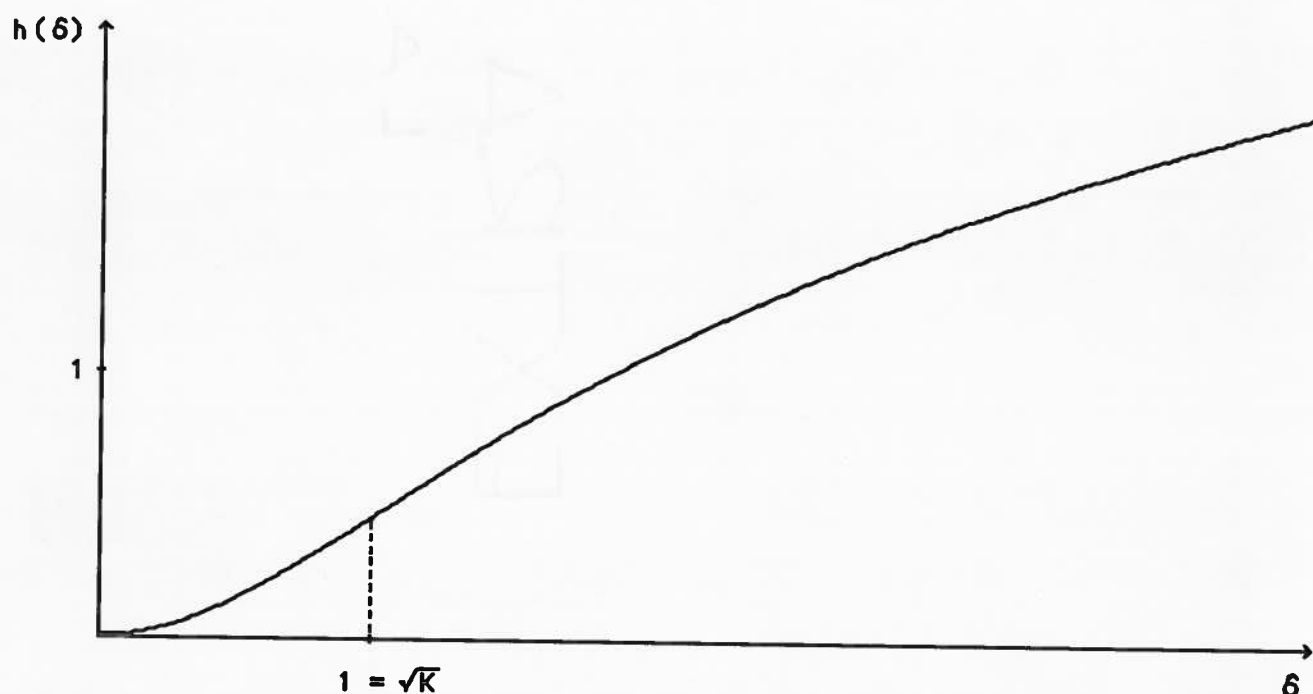


Fig. 2.5 — The t.c.f. $0.65 \ln(K + \delta^2)$, which generates circles as indifference lines ($K > 0$).
The inflexion point of the function occurs when $\delta = \sqrt{K}$. The number K is here equal to 1.

Function h is more complex in (e); see Fig. 2.5. It is here strictly convex on $[0, \sqrt{K}]$ and concave beyond; the marginal cost or disutility grows with distance in short trips, but decreases when the trip length exceeds $\sqrt{K}$; also, $h'(0) = 0$; and the centre of the disk Z_k lies at the right of k . In both cases (c) and (e), k may belong to its own market area or not. In item (d), and also somehow in (b), Z_k is more extended in the direction defined by j and k than in the perpendicular one (see Fig. 13.3 in Hanjoul et al.) In all items, Z_k is a convex set.

3. Variation of the transportation cost difference (and of y) along specific curves

In this chapter we review how the transportation cost difference Δh varies along some curves of the plane. Chapters 4 and 5 will make use of those 'microscopic' results to produce more elaborate statements, yet in their simple form they already yield immediate information about Z_j and Z_k in the following way. Suppose that a point i belongs to Z_k . If that point moves (thus creating a curve) so as to increase Δh , it obviously does not leave Z_k since Δh remains larger than Q . Intuitively, it moves inward relatively to Z_k (or away from Z_j). When Δh decreases, point i intuitively moves outward relatively to Z_k or toward Z_j . *Mutatis mutandis*, a similar property holds when $i \in Z_j$. Figures 3.1, 3.2 and 3.5 illustrate some points of this chapter by indicating, on curves of the appropriate type, the direction in which Δh increases.

Section 3.7 differs from the remaining of the chapter. It establishes concise formulæ of the variation of y on the indifference zone $Z_j \cap Z_k$, to be used in Chapters 4 and 5.

3.1. Move along circles centred on a point of $[jk]$ (Fig. 3.1)

In what follows, we denote by $C_p(r)$ the circle having p as centre and r as radius.

(3.1) For any point $p \in [jk]$ and any radius $r > 0$, Δh strictly decreases when φ_p increases, on the circle $C_p(r)$ above the x -axis.

If we move above the x -axis (i.e., inside $\mathcal{R} \times \mathcal{R}_+$) on a circle centred on p (i.e., with constant δ_p) so as to increase φ_p , it is obvious that δ_k increases strictly and δ_j decreases strictly when $p \in]jk[$. When $p = j$ (resp. k), δ_j (resp. δ_k) remains constant while δ_k increases (resp. δ_j decreases) strictly. As a consequence, Δh decreases strictly in every case.

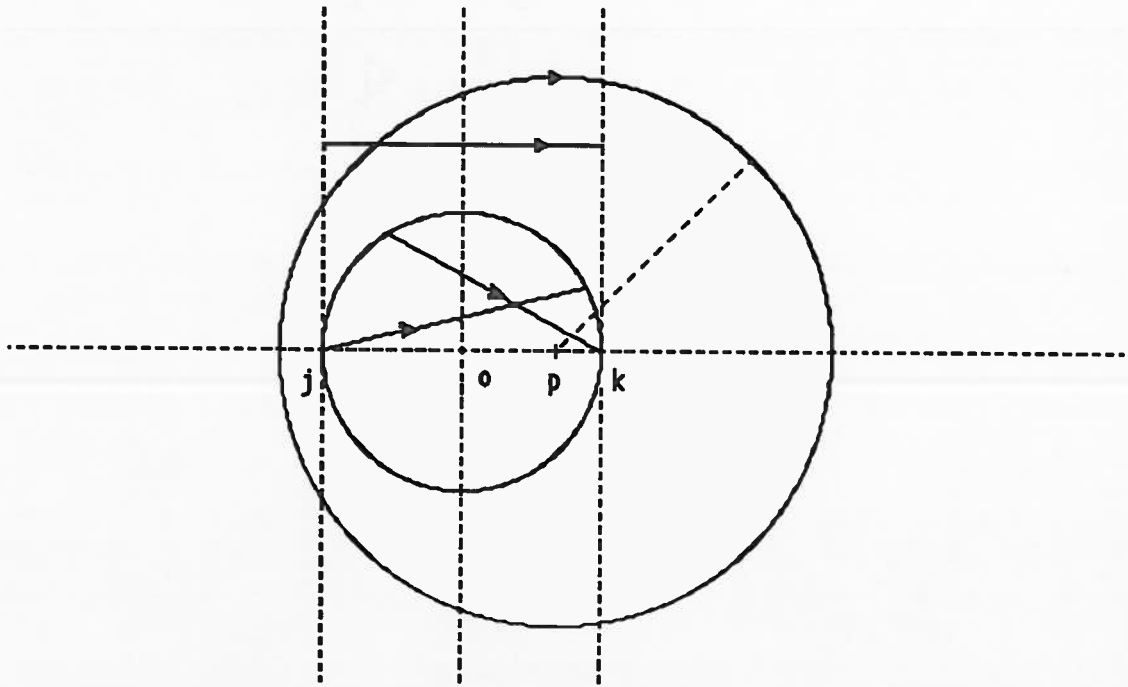


Fig. 3.1. — Direction of increase of Δh above the x -axis on several typical curves.

3.2. Radial moves from points of $[jk]$ (Fig. 3.1, 3.2, 3.3)

In the next proposition, we define $D(i, p)$ as the disk having i as centre and the circle of which passes through p .

- (3.2) (a) If function h is concave, Δh decreases when δ_k increases on the right of the y -axis on any straight line passing through k ;
- (b) If function h is convex and if p is some point of $[jo]$, Δh increases with δ_p on the right of the y -axis on any straight line passing through p ;
- (c) Inside the disk $D(o, k)$, Δh decreases when δ_k increases on any straight line passing through k , and Δh increases with δ_j on any straight line passing through j .

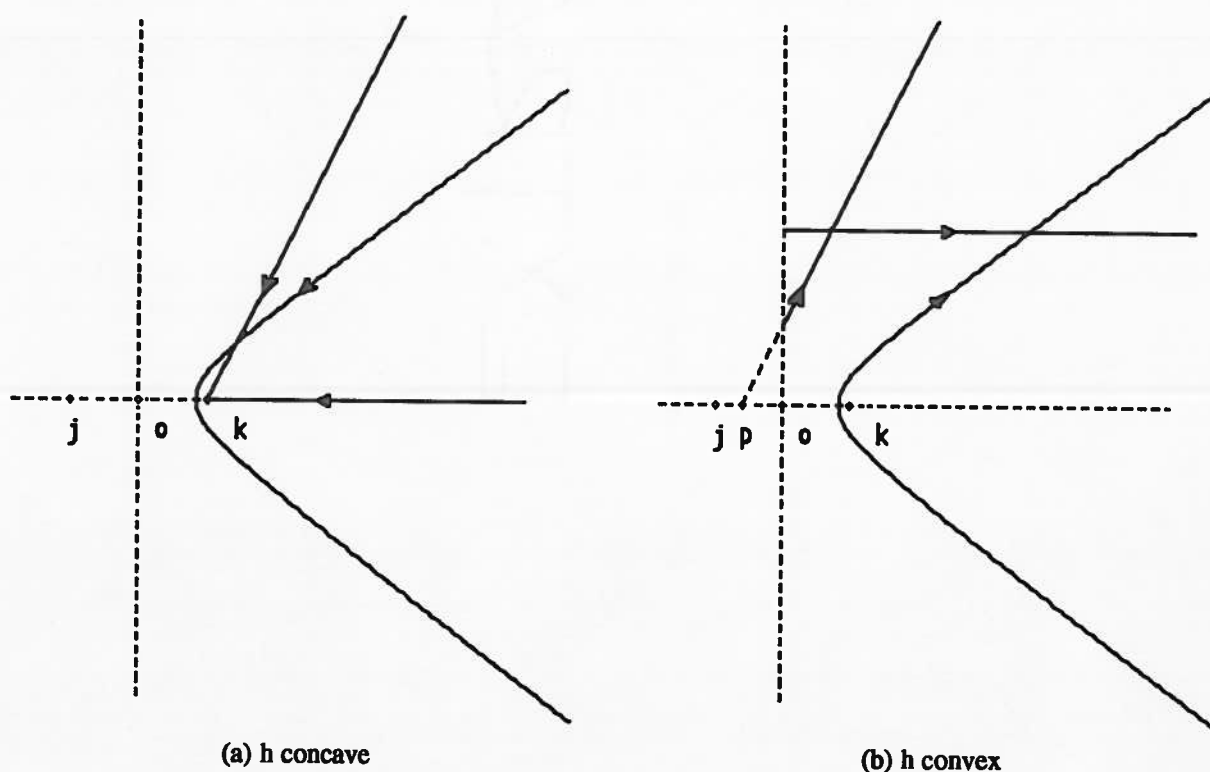


Fig. 3.2. — Direction of increase of Δh above the x -axis on some typical curves when h is concave or convex.

We give the proof of (c) about the straight lines passing through k ; the proof is similar concerning j . If a point i lies on such a line inside $D(o, k)$, the angle formed by j , i , and k is obtuse; so that if δ_k decreases, δ_j increases. Hence, Δh increases.

As to (a): consider some point i at the right of the y -axis (so that $\delta_j > \delta_k$), and some other point $i' \in]ik[$ (Fig. 3.3a). Letting $\lambda = \delta_{ii'}$, we thus have

$$\begin{aligned}
 \Delta h &= h(\delta_j) - h(\delta_k) \\
 &= h(\delta_j) - h(\delta_{i'k} + \lambda) \\
 &\leq h(\delta_{i'j} + \lambda) - h(\delta_{i'k} + \lambda) \quad (\text{triangle inequality}) \\
 &\leq h(\delta_{i'j}) - h(\delta_{i'k}) = \Delta h_{i'} \quad (h \text{ is concave, } \delta_{i'j} > \delta_{i'k}, \text{ and } \lambda > 0);
 \end{aligned}$$

which yields the proof.

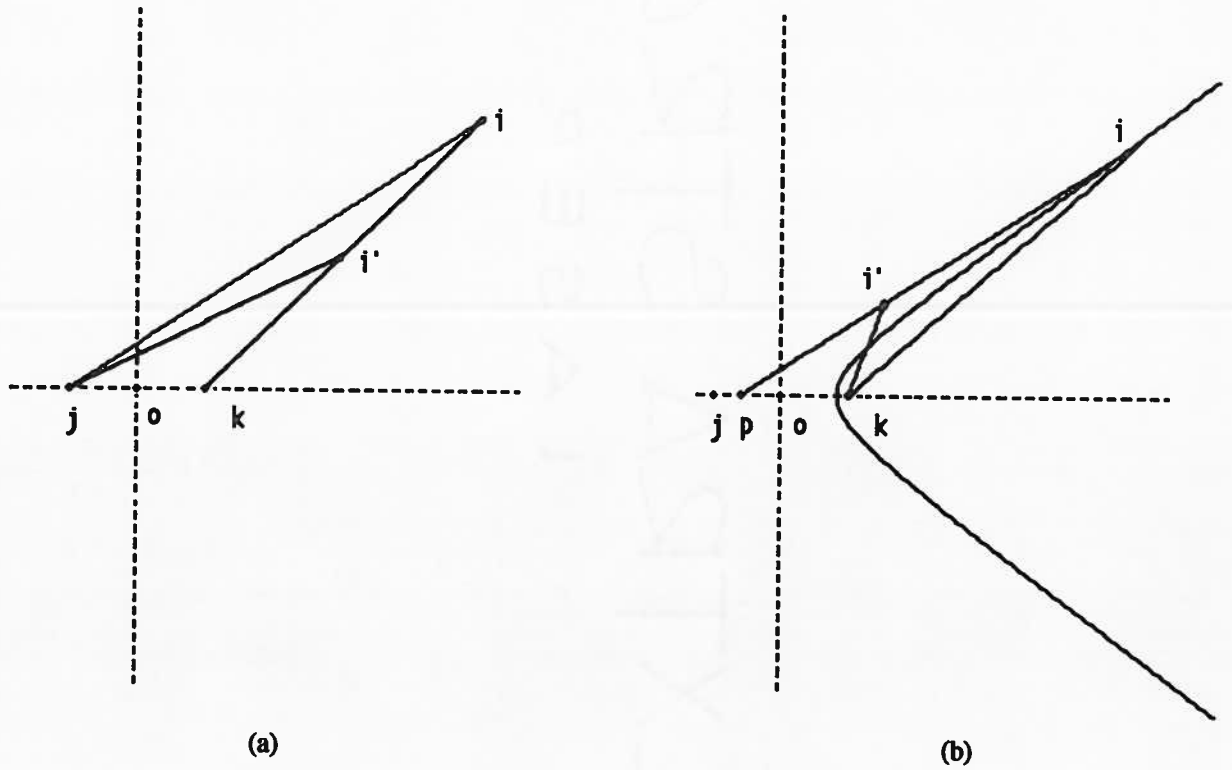


Fig. 3.3. — Proof of Prop. 3.2b and 3.2c

For point (b), we similarly take i on the right of the y -axis and $i' \in]ip[$ (Fig. 3.3b). First, as i is on the right of p , the projection of j on the straight line (ip) , on the one hand, and the points i and i' , on the other, are separated by the x -axis. Point i is thus farther than i' from j , so that $\delta_{i',j} < \delta_{ij}$. Second, if we consider the hyperbolic branch having j and k as focuses and passing through i , it is clear from the properties of hyperbolæ that $[pi[$ is entirely on the left of that branch; and that this means that $\delta_{i',j} - \delta_{i',k} < \delta_{ij} - \delta_{ik}$. Letting now $h(\delta) = h(0)$ when $\delta < 0$ so that h is convex on $\mathfrak{R}$ [$h(0) > -\infty$ since h is convex], and denoting $\delta_j - \delta_k$ as K and $\delta_j - \delta_{i',j}$ as λ (which is > 0 , as said above), we have

$$\begin{aligned} \Delta h_{i'} &= h(\delta_{i',j}) - h(\delta_{i',k}) \\ &\leq h(\delta_{i',j}) - h(\delta_{i',j} - K) && \text{as } \delta_{i',j} - \delta_{i',k} < K \\ &\leq h(\delta_{i',j} + \lambda) - h(\delta_{i',j} - K + \lambda) && \text{as } h \text{ is convex} \\ &&& \text{and } \lambda > 0 \end{aligned}$$

.3.5.

$$= h(\delta_j) - h(\delta_j - K) = h(\delta_j) - h(\delta_k) = \Delta h.$$

We thus have $\Delta h_i \leq \Delta h$, which yields the proof. Q.E.D.

Noteworthily, the proof of item (a) is very general and the statement is easy to extend to any transportation space. The same is true for item (b) when $p = j$. The statement then holding is somehow the symmetric of item (a) for convex t.c.f.'s. A proof of that restricted property exists that is closely similar to that of item (a) and shows its extendability to any transportation space.

3.3. Move along horizontal straight lines (fig. 3.1, 3.2)

(3.3) (a) Δh increases strictly with x on any horizontal line in the strip $[-\delta_{ok}, \delta_{ok}] \times \mathbb{R}$;

(b) If function h is convex, Δh increases with x on any horizontal line on the right of the y -axis;

(c) If function h is concave [resp. convex], Δh decreases [resp. increases] when x increases on the x -axis at the right on k , i.e., on $]\delta_{ok}, +\infty[\times \{0\}$.

Item (a) is obvious: when i lies in the mentioned strip and x increases while the point remains in the strip, δ_j increases strictly but δ_k decreases strictly, so that Δh increases strictly.

Now for item (b). According to (2.3), we know that

$$\begin{aligned} (\partial \delta_j / \partial x)_y &= (x + \delta_{ok}) / \delta_j = \cos(\varphi_j) \\ (\partial \delta_k / \partial x)_y &= (x - \delta_{ok}) / \delta_k = \cos(\varphi_k) . \end{aligned} \tag{3.1}$$

It is well-known that when a function is concave or convex on an open interval, it is also left- and right-derivable on it. Let us denote the right derivative of h on $]0, \infty[$ as h'_+ . Let us fix y and increase x , starting from a

point i for which $x > \delta_{0k}$, since the case of $x \in [-\delta_{0k}, \delta_{0k}]$ has already been considered in item (a). Both δ_j and δ_k increase; thus, with obvious notation,

$$\begin{aligned}
 (\partial \Delta h / \partial x)_{y+} &= h'_{+j} \cos(\varphi_j) - h'_{+k} \cos(\varphi_k) \\
 &= \cos(\varphi_j) (h'_{+j} - h'_{+k}) + h'_{+k} [\cos(\varphi_j) - \cos(\varphi_k)] \\
 &> 0 \text{ if } h \text{ is convex, as } \cos(\varphi_j) > 0 \text{ and} \\
 &0 \leq |\varphi_j| \leq |\varphi_k| \leq \pi.
 \end{aligned} \tag{3.2}$$

As to (c), on the x -axis at the right of k , we have $\Delta h = h(\delta_k + Q) - h(\delta_k)$. It is well-known that if Q is constant and positive, Δh must then be increasing w.r.t. δ_k when h is convex, and decreasing when h is concave. **Q.E.D.**

3.4. Crossing between indifference lines and other hypercircles

In this section we study the variation of Δh along hypercircles associated with some auxiliary t.c.f., $h^\sim$.

(3.4) *If $h(\delta)$ is a convex (resp. concave) expression of $h^\sim(\delta)$, Δh increases (resp. decreases) when δ_k increases along the hypercircles associated with $h^\sim$ at the right of the y -axis.*

To prove Prop. 3.4, we just have to rewrite Δh as follows:

$$\Delta h = h_j - h_k = T(h^\sim_j) - T(h^\sim_k).$$

The relation T , equal to $h \circ h^\sim$, is indeed a function: $h^\sim$ is a function since $h^\sim$ is injective, being strictly monotonous. Hence it is clear that a sufficient condition for Δh to increase with δ_k (or, equivalently, with $h^\sim_k$) along the hypercircle represented by the equality $h^\sim_j = h^\sim_k + Q^\sim$, where $Q^\sim > 0$, is that T be convex; in other terms, that the expression $h(\delta)$ be concave w.r.t. $h^\sim(\delta)$. (The remainder of the proof is obvious). **Q.E.D.**

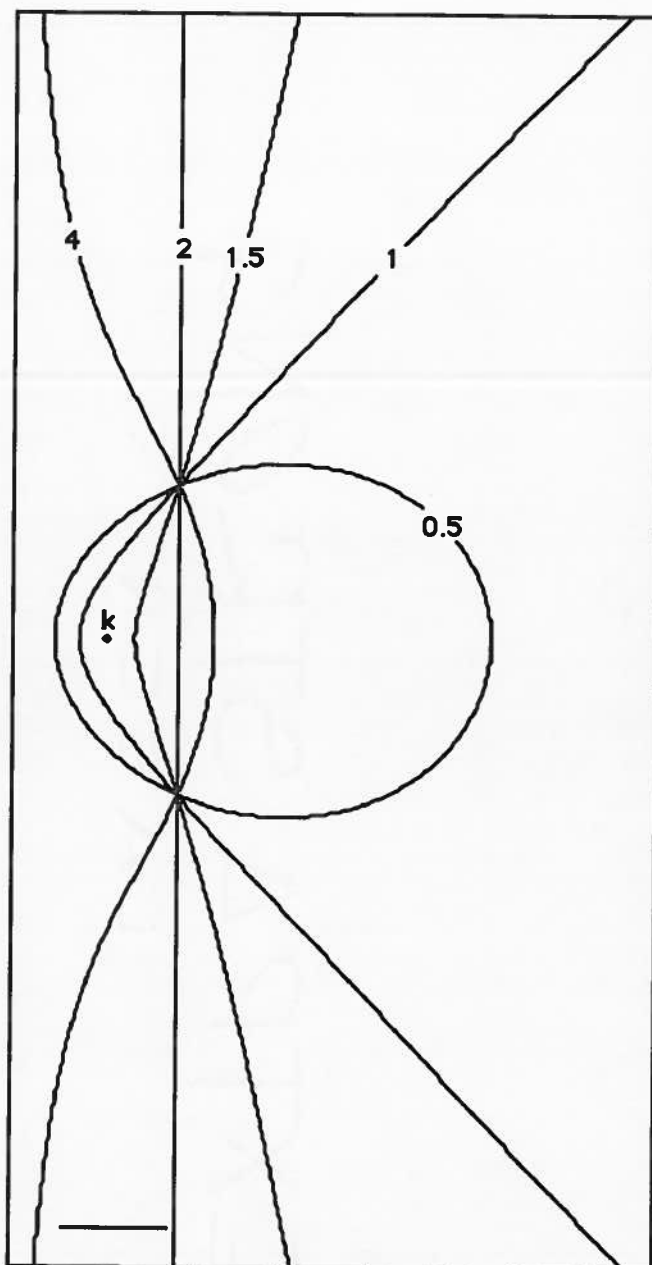


Fig. 3.4. — Market boundaries generated by the power t.c.f. δ^a for indicated values of a when $\delta_{jk} = 1.75$, Q being chosen so as to let the indifference line pass through the point $(1.625, 1.375)$.

Two remarks are to be made concerning Prop. 3.4: 1° according to Proposition 2.4, δ_k may be replaced by δ_j in that statement; or by δ_p , for any $p \in [jk]$; 2° when $h \cdot h^\wedge$ is *strictly* convex or concave, the monotonicity of Δh along the hypercircles is also strict.

To illustrate Prop. 3.4, assume for instance that $h(\delta) = \delta^a$. If $h^\sim(\delta) = \delta^{\tilde{a}}$, we have $T[h^\sim(\delta)] = \delta^a = (\delta^{\tilde{a}})^{a/\tilde{a}}$. Suppose that $0 < a < \tilde{a}$; T is

then a concave function. If p is a point on $Z_j \cap Z_k$ and if, starting from p , we move on the curve $Z_j \cap Z_k$ passing through p , Prop. 3.4 and 2.4 show that we remain in Z_j when we move away from $[jk]$, or in Z_k when we move towards $[jk]$. This may be observed on Fig. 3.4, where $Z_j \cap Z_k$ is constrained to pass through a point p , with various values of the exponent a . The fact that $Z_j \cap Z_k$ moves anticlockwise when the exponent a increases is easily proved on basis of the expression of $(\partial y / \partial x)_{\Delta h}$; see Section 3.7.

We may also combine Propositions 2.7 and 3.4 (see Fig. 3.2 and 3.5) :

- (3.5) (a) When $h \circ \exp$ is concave (resp. convex), Δh decreases (resp. increases) when δ_k increases on any Apollonian circle relative to j and k , at the right of the y -axis;
- (b) When h is concave (resp. convex), Δh decreases (resp. increases) when δ_k increases on any hyperbola having j and k as foci, at the right of the y -axis;
- (c) When $h \circ \sqrt{\cdot}$ is concave (resp. convex), Δh decreases (resp. increases) when y increases on any vertical line, at the right of the y -axis.

(See the note that follows Proposition 3.4). Notice also that Prop. 3.3c may be viewed as an application of Prop. 3.5b, as $[\delta_{ok}, +\infty[\times \mathfrak{R}$ is the right-hand branch of the degenerate hyperbola described by the equality $\delta_j - \delta_k = \delta_{jk}$.

As to the concavity or convexity of $h \circ \exp$ and $h \circ \sqrt{\cdot}$, they are immediately known in some important cases :

- (3.6) (a) When h is convex, or when $h(\delta) = \delta^a \forall \delta \geq 0$ with $a > 0$,
then $h \circ \exp$ is convex ;
- (b) When h is concave, or when $h(\delta) = \delta^a \forall \delta \geq 0$ with $0 < a \leq 2$,
then $h \circ \sqrt{\cdot}$ is concave ;
- (c) When $h(\delta) = \delta^a \forall \delta \geq 0$ with $2 \leq a$,

then $h \circ \sqrt{\cdot}$ is convex.

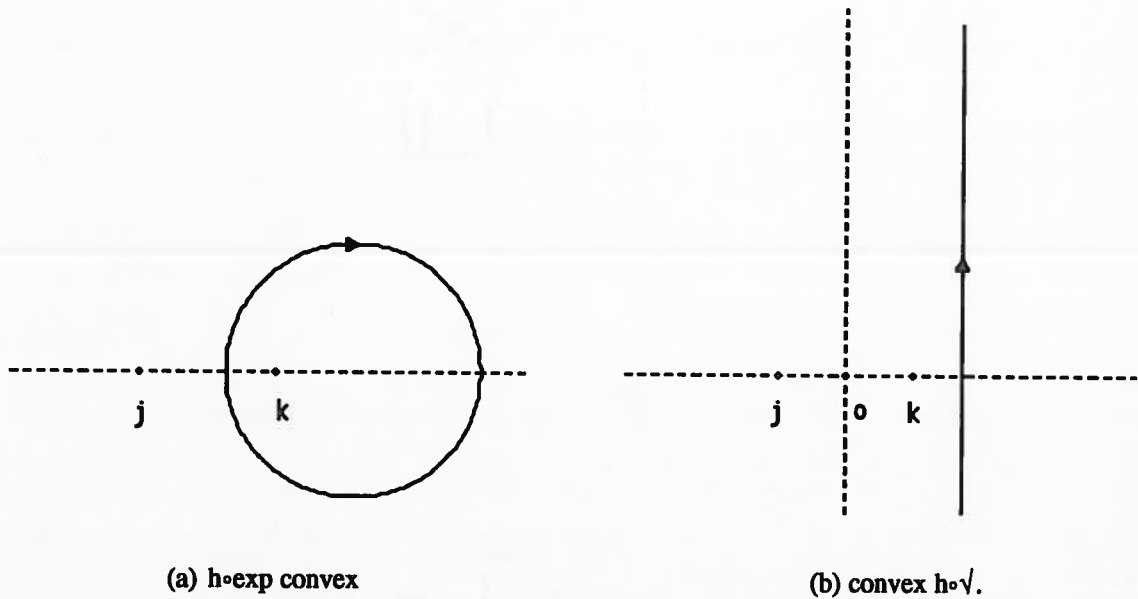


Fig. 3.5. — Direction of increase of Δh above the x -axis on hypercircles associated with the t.c.f.'s $\ln(\delta)$ and δ^2 .

Except when $h(\delta) = \delta^a$, the proof is that any increasing convex (resp. concave) function of a convex (resp. concave) function is convex (resp. concave). Reciprocally, we have too

- (3.7) (a) When $h \circ \exp$ is concave, h is concave;
 (b) When $h \circ \sqrt{\cdot}$ is convex, h is convex.

3.5. Relative position of the indifference line and some curves associated with another t.c.f.

Considering again an auxiliary t.c.f. h^- , we define, for any point $i \in Z_j \cap Z_k$, the area

$$G_i = \{p; h^-_{kp} \leq \alpha_i h^-_{jp} + \zeta_i\}.$$

The parameters α_i and ζ_i are chosen so that the curve E_i defined by replacing ' $\leq$ ' by '=' in that inequality is tangent to $Z_j \cap Z_k$ in i . Which means, if h

and h^- are derivable on $]0, \infty[$, that i belongs to that curve and that $\alpha_i = \Phi'(h^-_j)$, supposing that we have put the equality defining $Z_j \cap Z_k$ in the form: $h^-_k = \Phi(h^-_j)$. As the inverse relations h^+ and h^- are functions since h and h^- are strictly monotonous and thus injective, we have indeed, on $Z_j \cap Z_k$,

$$\begin{aligned} h^-_k &= h^- [h^+ (h_k)] = h^- [h^+ (h_j - Q)] \\ &= (h^- \circ h^+) [(h \circ h^-) (h^-_j) - Q] = T^+ [T (h^-_j) - Q] \\ &= \Phi (h^-_j) , \end{aligned}$$

where we have set $T = h \circ h^-$ as in Section 3.4, and

$$\Phi = T^+ \circ (T - Q) . \quad (3.3)$$

As to the derivatives of T , T^+ , and Φ , we have

$$\begin{aligned} T' &= (h \circ h^-)' = (h' \circ h^-) / (h'^- \circ h^-) \\ &= (h' / h'^-) \circ h^- \end{aligned} \quad (3.4a)$$

$$T^{+'} = (h'^- / h') \circ h^+ \quad (3.4b)$$

$$\Phi' = [T^{+'} \circ (T - Q)] \cdot T' = T' / T' \circ T^{+'} \circ (T - Q) = T' / T' \circ \Phi \quad (3.4c)$$

and we find

$$\begin{aligned} \alpha_i &= \Phi'(h^-_j) = T' (h^-_j) / T' (h^-_k) = h'^-_k h'_j / h'^-_j h'_k \\ \zeta_i &= (h^-_k h'_k / h'^-_k - h^-_j h'_j / h'^-_j) h'^-_k / h'_k . \end{aligned} \quad (3.5)$$

On the other hand, if there is a point ℓ as defined in Chapter 2, we also consider the area

$$G_{\ell r} = \{p; h^-_{kp} \leq \alpha_{\ell r} h^-_{jp} + \zeta_{\ell r}\},$$

where $\alpha_{\ell r}$ and $\zeta_{\ell r}$ are chosen so that the curve $E_{\ell r}$ defined by the equality corresponding to that inequality passes through the points r and ℓ .

The inequalities defining G_i and G_{ℓ_r} are similar to (1.2), and the Chapter 13 of Hanjoul et al. may be applied to them. We have the following property.

(3.8) If $\ln \circ (h^{-1}/h') \circ h^{\wedge}$ is concave (resp. convex) on $h > 0$, $\infty[$, $G_i \subseteq Z_k$ (resp. $Z_k \subseteq G_i$) for any $i \in Z_j \cap Z_k$;

if moreover the point ℓ exists, $Z_k \subseteq G_{\ell_r}$ (resp. $G_{\ell_r} \subseteq Z_k$).

First of all, let us notice that, because of the continuity (*lato sensu*) of h in 0 , we have not to bother about what happens when $\delta = 0$, at the extremity of $]0, \infty[$, since this would only concern the belonging or not of centre k to G_i or G_{ℓ_r} . The expressions defining those areas are indeed continuous *lato sensu* in k w.r.t. the coordinates (δ_j, δ_k) . Not to overlast the writing, we consider in the remainder of this proof that the definition set of the functions h and h^{-1} is $]0, \infty[$ (instead of $[0, \infty[$); the definition sets of T , $T^{\wedge}$, and Φ are thus also affected (the definition set or domain of a function is the set on which the function takes real values). We so make the proof easier to explain since we avoid speaking of the continuity *lato sensu* of h in 0 .

Seen in the generalized bipolar coordinates (h^{-1}_j, h^{-1}_k) , the real region 'under' $Z_j \cap Z_k$ is contained in Z_k according to Prop. 3.1 and its proof; whereas, similarly, the real region under E_{ℓ_r} is in G_{ℓ_r} . 'Real' means here that the coordinates (h^{-1}_j, h^{-1}_k) correspond to some point of the plane. A couple (h^{-1}_j, h^{-1}_k) is indeed 'under' another if it has the same δ_j but a smaller δ_k . Moreover, the curves E_i and E_{ℓ_r} (if any E_{ℓ_r} ; i.e., if Z_k is bounded but not empty, as will be shown in Chapter 4) appear as straight; more precisely, as the real parts of straight lines. Consequently, a sufficient condition for Z_k to contain G_i and to be contained in G_{ℓ_r} is that $Z_j \cap Z_k$ appears in those coordinates as the graph, or some part of the graph, of a convex function. Otherwise said: that the function Φ defined above is convex; which means that Φ' is increasing, provided that $\text{dom}(\Phi)$ (the definition set of Φ) is an interval and that the derivative Φ' is real-valued on $\text{dom}^\circ(\Phi)$ (the interior of that interval).

Let us first see if those conditions hold about $\text{dom}(\Phi)$ and the derivability of Φ . As $\Phi^{\wedge}$, equal to $h^{-1} \circ h^{\wedge} \circ (1 \otimes \text{id}) \circ h \circ h^{\wedge}$, is a concatenation of functions, each of which has a real interval as definition set $]0, \infty[$,

$h <]0, \infty[$, $\mathfrak{X}$, $]0, \infty[$, and $h^{-} <]0, \infty[$, respectively) and is continuous on it, the image of any real interval through $\Phi^{\wedge}$ is an interval; and such is consequently $\text{dom}(\Phi)$, being of course equal to $\Phi^{\wedge}(\mathfrak{X})$. As to the derivative Φ' , we see from (3.3) that both $\text{dom}(\Phi)$ and $\Phi^{\wedge}(\mathfrak{X})$ are contained in $\text{dom}(T)$; hence, (3.4c) implies that, if T is derivable on $\text{dom}^{\circ}(T)$ and if T' does nowhere take 0 as value, Φ is derivable on $\text{dom}^{\circ}(\Phi)$. And, concerning T' , we similarly see from the definition of T that $\text{dom}(T) = T^{\wedge}(\mathfrak{X}) = h^{-} <]0, \infty[$, which is an interval; and thus, from (3.4a), that T is derivable on the interior of that interval if h and h^{-} are derivable on $]0, \infty[$ and if h^{-} does nowhere take 0 as value. Those conditions on h and h^{-} are guaranteed by our hypothesis that $\ln \circ (h^{-}/h') \circ h^{\wedge}$ is concave, and thus real-valued, on $h <]0, \infty[$; Φ is thus derivable on $\text{dom}^{\circ}(\Phi)$ as required.

Let us now study the possible increasingness of Φ' . As T' and Φ' are positive, we have, according to (3.4c),

$$\ln \circ \Phi' = \ln \circ T' - \ln \circ T' \circ \Phi.$$

Hence a sufficient condition for Φ' to be increasing is that $\ln \circ T'$ is convex. [Which means the concavity of the expression $\ln[h^{-}(\delta)/h'(\delta)]$ w.r.t. $h(\delta)$]. Similarly, a sufficient condition for Φ' to be decreasing is that $\ln \circ T'$ is concave. **Q.E.D.**

When the functions h and h^{-} are twice derivable, the condition of Prop. 3.8 may be expressed in alternative ways :

(3.9) Assume that, on $]0, \infty[$, 1° h'' and $h^{-''}$ are real-valued, and 2° h^{-}' never takes 0 as value. Then the following three properties are equivalent:

- (i) $\ln \circ (h^{-}'/h') \circ h^{\wedge}$ is concave (resp. convex) on $h <]0, \infty[$,
- (ii) $(h^{-}'/h') \circ h^{\wedge}$ is concave (resp. convex) on $h^{-} <]0, \infty[$,
- (iii) $(h^{-}'/h')' / h^{-}'$, i.e. $(h^{-''}/h^{-}' - h''/h') / h'$, is decreasing (resp. increasing) on $]0, \infty[$.

For this proof, we restrict the domains of the concerned functions as in the proof of Prop. 3.8. As to (i), we have

$$[\ln \circ (h^{-}'/h') \circ h^{\wedge}]'$$

$$\begin{aligned}
 &= [1/(h'/h') \circ h^{\wedge}] \cdot [(h'/h')' \circ h^{\wedge}] \cdot [1/(h' \circ h^{\wedge})] \\
 &= [(h'/h') \cdot (h'/h')'/h'] \circ h^{\wedge} \\
 &= [(h'/h')'/h'] \circ h^{\wedge} .
 \end{aligned} \tag{3.6}$$

In which we have :

$$\begin{aligned}
 (h'/h')'/h' &= (h'' \cdot h' - h' \cdot h'') / h'^2 \cdot h' \\
 &= (h''/h' - h'/h'') / h' .
 \end{aligned}$$

This shows that $(h'/h')'/h'$ is real-valued on $]0, \infty[$ under our hypotheses. Moreover, as $h^{\wedge}$ is continuous and increasing on $h > 0$, equality (3.6) shows that the increasingness (resp. decreasingness) of $(h'/h')'/h'$ implies that of $[(h'/h')'/h'] \circ h^{\wedge}$. Similarly, as h is continuous and increasing on $]0, \infty[$ and $[(h'/h')'/h'] \circ h^{\wedge}$ is real-valued on $h > 0$, the increasingness (resp. decreasingness) of $[(h'/h')'/h'] \circ h^{\wedge}$ implies that of $[(h'/h')'/h'] \circ h^{\wedge} \circ h$, i.e., of $(h'/h')'/h'$. The equivalence of (i) and (iii) is thus established.

Regarding (ii), we have

$$\begin{aligned}
 [(h'/h') \circ h^{\wedge}]' &= [(h'/h')' \circ h^{\wedge}] / (h' \circ h^{\wedge}) \\
 &= [(h'/h')'/h'] \circ h^{\wedge} ,
 \end{aligned}$$

which resembles (3.6), and we conclude, similarly, that (ii) and (iii) are equivalent. **Q.E.D.**

Of course, the condition (iii), although less general than (i), will be used most of the time; the condition (ii) will appear useful in Prop. 3.10. [Hanjoul et al. (1988) only mention the condition (ii), of which they give another proof]. Let us for instance apply (iii) to situations in which both h and $h^{\wedge}$ are monomial functions: $h(\delta) = \delta^a$ and $h^{\wedge}(\delta) = \delta^{\tilde{a}}$, with $a > 0$ and $\tilde{a} > 0$. We have here

$$\begin{aligned}
 [(h'/h')'/h'](\delta) &= \{(\tilde{a}/a)[(\delta)^{\tilde{a}-a}]'/\tilde{a}(\delta)^{\tilde{a}-1}\}(\delta) \\
 &= (1/a)(\tilde{a} - a)\delta^{-a} .
 \end{aligned}$$

That expression is decreasing $[G_i \subseteq Z_k \subseteq G_{lr} \text{ (if any)}]$ when $0 < a \leq \tilde{a}$; and increasing $[G_{lr} \text{ (if any)} \subseteq Z_k \subseteq G_i]$ when $a \geq \tilde{a}$.

Two at least of those cases yield interesting results. When $\tilde{a} = 1$, i.e., when $h^\sim$ is the identity function, the associated curves E_i and E_{lr} are Cartesian ovals. The second case occurs when $\tilde{a} = 2$. The inequality defining G_i then writes

$$(\alpha_i - 1) [(x_p - \delta_{ok} \frac{1+\alpha_i}{1-\alpha_i})^2 + y_p^2] \geq 4\delta_{ok}^2 \frac{\alpha_i}{\alpha_i-1} - \zeta_i ; \quad (3.7)$$

hence, E_i is a circle and its centre, $[\delta_{ok} (1+\alpha_i)/(1-\alpha_i), 0]$, is on the x-axis; we shall come back to that property in Sections 4.4 and 5.5. As this has remarkable consequences, Prop. 3.9 indicates that the functions $\ln^\circ[h^\sim/(h' \cdot h^\sim)]$, $\sqrt{\cdot}/(h' \cdot \sqrt{\cdot})$, and $(\ln^\circ/h')'/h'$ are particularly important in the study of the shape of market areas.

To end this section, notice that Prop. 3.8 finds its place in the present chapter, as far as the curves E_i are concerned, since it may be expressed in terms of extrema of Δh on those curves. On the other hand, Prop. 3.8 differs from all the preceding ones of this chapter in that its conditions are not additive. For those statements indeed, if the conditions are verified for a set of transportation cost functions, they remain so for any positive-coefficient linear combination of those functions. This is not true for Prop. 3.8. However, Prop. 3.8 is partially additive, as will now be seen.

3.6. On parallel association

Given any real numbers a and b , we set this definition:

$$\text{if } a + b \neq 0 : \quad a // b = ab / (a + b) .$$

We call the number $a//b$ the *parallel association* of a and b . The concept is borrowed from the domain of electricity, where that formula describes the parallel association of two impedances. In fact, if $a \neq 0 \neq b$ and $a + b \neq 0$, we have $a // b = 1/(1/a + 1/b)$. Another interesting property: parallel

association is associative. Given indeed any three numbers a , b , and c , the quantities $a/(b/c)$ and $(a/b)/c$, when they may be computed, are equal and may be written as $a/b/c$.

LEMMA 3.1. *Let a , b , c , and d be real numbers having these properties: $a \geq c$, $b \geq d$, and $c + d > 0$.*

Then $a // b \geq c // d$.

Three situations are possible concerning the position of 0 with respect to a and c : $a \geq c > 0$, $a \geq 0 \geq c$, and $0 > a \geq c$. On the other hand, we have three situations characterized by these inequalities: $b \geq d > 0$, $b \geq 0 \geq d$, and $0 > b \geq d$. As $c + d > 0$, c and d cannot be ≤ 0 at the same time, so that those situations may be combined into only five situations : 1° $a \geq c > 0$ and $b \geq d > 0$; 2° $a \geq 0 \geq c$ and $b \geq d > 0$; 3° $0 > a \geq c$ and $b \geq d > 0$; 4° $a \geq c > 0$ and $b \geq 0 \geq d$; 5° $a \geq c > 0$ and $0 > b \geq d$. In situations 2 and 4, we have $a/b \geq 0 \geq c/d$ and the lemma holds true (notice that $a+b > c+d > 0$ in any case).

In the remaining situations, a and c bear the same sign; so do b and d , and none of the numbers a , b , c , and d is zero. Hence, the inequalities $a \geq c$ and $b \geq d$ imply in those cases that $1/a \leq 1/c$ and $1/b \leq 1/d$; from which $1/a + 1/b \leq 1/c + 1/d$; i.e., $1/(a/b) \leq 1/(c/d)$. In situation 1°, a/b and c/d are > 0 , and in situations 3° and 5°, a/b and c/d are < 0 . In each of the situations 1°, 3°, and 5°, a/b and c/d thus bear the same sign; hence, the inequality $1/(a/b) \leq 1/(c/d)$ results in: $a/b \geq c/d$. The lemma is thus verified in every case. **Q.E.D.**

We are now in a position to establish a second lemma :

LEMMA 3.2. *Let f and g be two concave functions having a same definition set.*

If $f + g > 0$, $f // g$ is concave.

The concavity of f means that for every a and b of the definition set E , and for every $x \in [a, b] \cap E$, $f(x) \geq \phi(x)$, in which formula the function ϕ is the linear function for which $\phi(y) = f(y)$ and $\phi(z) = f(z)$. The function γ

being defined similarly with regard to the function g and the same numbers a and b , we also have $g(x) \geq \gamma(x) \forall x \in [a, b] \cap E$. On the other hand, as $f + g > 0$, it is clear that $\phi + \gamma$ is > 0 on $[a, b]$. Lemma 3.1 may thus be applied, so that $(f//g)(x) \geq (\phi//\gamma)(x) \forall x \in [a, b] \cap E$.

The definition of $//$ shows that $\phi//\gamma$ takes real values over $[a, b]$ since $\phi + \gamma$ is > 0 on that set. Let us compute the first two derivatives of that function; we denote as α and β the first derivatives of ϕ and γ , respectively.

$$\begin{aligned}
 (\phi//\gamma)' &= [\phi\gamma / (\phi + \gamma)]' \\
 &= [(\alpha\gamma + \beta\phi)(\phi + \gamma) - \phi\gamma(\alpha + \beta)] / (\phi + \gamma)^2 \\
 &= (\alpha\gamma\phi + \alpha\gamma^2 + \beta\phi^2 + \beta\gamma\phi - \alpha\gamma\phi - \beta\gamma\phi) / (\phi + \gamma)^2 \\
 &= (\alpha\gamma^2 + \beta\phi^2) / (\phi + \gamma)^2 \\
 (\phi//\gamma)'' &= [(2\alpha\beta\gamma + 2\alpha\beta\phi)(\phi + \gamma)^2 - (\alpha\gamma^2 + \beta\phi^2) 2(\phi + \gamma)(\alpha + \beta)] \\
 &\quad / (\phi + \gamma)^4 \\
 &= [2\alpha\beta(\phi + \gamma)^3 - (\alpha\gamma^2 + \beta\phi^2) 2(\phi + \gamma)(\alpha + \beta)] \\
 &\quad / (\phi + \gamma)^4 \\
 &= 2[\alpha\beta(\phi + \gamma)^2 - (\alpha + \beta)(\alpha\gamma^2 + \beta\phi^2)] \\
 &\quad / (\phi + \gamma)^3 \\
 &= 2(\alpha\beta\phi^2 + 2\alpha\beta\gamma\phi + \alpha\beta\gamma^2 - \alpha^2\gamma^2 - \alpha\beta\phi^2 - \alpha\beta\gamma^2 - \beta^2\phi^2) \\
 &\quad / (\phi + \gamma)^3 \\
 &= -2(\alpha^2\gamma^2 + \beta^2\phi^2 - 2\alpha\beta\gamma\phi) / (\phi + \gamma)^3 \\
 &= -2(\alpha\gamma - \beta\phi)^2 / (\phi + \gamma)^3.
 \end{aligned}$$

The function $\phi//\gamma$ is thus concave over $[a, b]$. We have thus $(\phi//\gamma)(x) \geq \psi(x) \forall x \in [a, b]$, where ψ is the linear function defined by the equalities $\psi(a) = (\phi//\gamma)(a)$ and $\psi(b) = (\phi//\gamma)(b)$. However, $\forall x \in [a, b] \cap E$, we have seen that $(f//g)(x) \geq (\phi//\gamma)(x)$, so that $(f//g)(x) \geq \psi(x)$; and $(f//g)(a) = (\phi//\gamma)(a)$ and $(f//g)(b) = (\phi//\gamma)(b)$, so that $(f//g)(a) = \psi(a)$ and $(f//g)(b) = \psi(b)$. It is thus clear that $f//g$ is concave. Q.E.D.

Notice that the proof also shows that if the condition ' $f + g > 0$ ' is not met, it is possible to find concave functions having a nonconcave parallel association: this is patent in the case of linear functions.

Lemma 3.2 enables us to extend the practicability of Prop. 3.8, on the basis of Prop. 3.9. We may indeed state the following:

(3.10) Let h be a positive-coefficient linear combination of transportation cost functions h_v , $v \in \{1, \dots, n\}$. If the following conditions are satisfied for any $v \in \{1, \dots, n\}$: 1° on $]0, \infty[$, h_v'' and h_v'' are real-valued and h_v' does not take 0 as value, and 2° $(h_v'/h_v) \cdot h_v'$ is concave on $h_v' <]0, \infty[$,

then $(h'/h) \cdot h'$ is also concave on $h' <]0, \infty[$.

The operator '//' being associative, $(h'/h) \cdot h_v'$ is the parallel association of the functions $f_v = (h_v'/h_v) \cdot h_v'$, with $v \in \{1, \dots, n\}$. If restricted to $h' <]0, \infty[$, the functions f_1 and f_2 satisfy the requirements of Lemma 3.1: they are concave, and f_1 and f_2 are > 0 (since h_v' , h_v' , and h_v' are > 0 on $]0, \infty[$) so that $f_1 + f_2 > 0$. Hence, $f_1//f_2$ is concave on $h' <]0, \infty[$. As $f_1//f_2$ and f_3 also satisfy those requirements, $f_1//f_2//f_3$ is also concave on $h' <]0, \infty[$. And so on; $(h'/h) \cdot h'$ is thus concave on $h' <]0, \infty[$. Q.E.D.

Prop. 3.10, combined with Prop. 3.8 and 3.9, ensures for instance that if the t.c.f. is defined as $\delta + \sqrt{\delta}$ and if h' is the identity function, the area G_1 , limited by a Cartesian oval, is contained in Z_k for any $i \in Z_j \cap Z_k$.

On the other hand, the definition of the operator '//' incidentally provides us with an original way to build new t.c.f.'s:

(3.11) If h and h' have the same sign everywhere, or if there is some δ^* for which $h(\delta^*) = h'(\delta^*) = 0$ and the limit of $(h/h')(\delta)$ when δ tends toward δ^* is 0, then the function equal to h/h' on $[0, \infty[-\{\delta^*\})$ and taking 0 as value in δ^* is a t.c.f.

[Here again, for continuity reasons, we have not to consider $h(0)$]. Assume first that h and $h^\sim$ never equal 0. As h and $h^\sim$ are real strictly increasing functions, $1/h$ and $1/h^\sim$ are strictly decreasing (be they positive or negative) and have the same sign, so that $1/h + 1/h^\sim$ is strictly decreasing; hence, $h/h^\sim$ is strictly increasing and may thus be considered as a t.c.f. Now if $h(\delta^*) = h^\sim(\delta^*) = 0$ for some δ^* (necessarily unique since both functions are strictly increasing), $h/h^\sim$ is strictly increasing on $[0, \delta^*[$ and on $] \delta^*, \infty[$. (The remainder of the proof is obvious). **Q.E.D.**

For instance, the expressions $\delta/\sqrt{\delta}(\delta+\sqrt{\delta})$ and $-1/(\delta+\delta^2)$ define t.c.f.'s, as being equal to $\delta/\sqrt{\delta}$ and $(-1/\delta)/(-1/\delta^2)$, respectively. They are concave, moreover (the former according to Lemma 3.1). A more complex example is given by the t.c.f.'s defined by the expressions $\delta-1$ and $\ln(\delta)$: both are < 0 for $\delta < 1$, and > 0 for $\delta > 1$, and it is easily computed by de l'Hospital's rule that the limit of $h/h^\sim$ when δ tends toward 1 is 0; hence, the expression $(\delta-1)/\ln(\delta)$ and its continuous prolongation in 1 define a t.c.f.

3.7. Variation of y with x along the indifference line

This section provides information about the slope of $Z_j \cap Z_k$. It establishes practicable formulæ of the first and second derivatives of y w.r.t. x on $Z_j \cap Z_k$. The second derivative will be used in the proof of Prop. 5.6 to show the convexity of market area Z_k in some cases. The first derivative leads of course to the second, and also to some properties which we shall mention although they may also be seen as corollaries of subsequent ones of Chapter 5.

A simple way to arrive at an expression of the first derivative is to consider $\partial y^2/\partial x$, as

$$\partial y^2/\partial x = 2y \partial y/\partial x . \quad (3.8)$$

The second equality of (2.3) yields

$$\partial y^2/\partial x = \partial \delta_k^2/\partial x - 2 (x - \delta_{ok}) . \quad (3.9)$$

On $Z_j \cap Z_k$, we have

.3.19.

$$(h \circ \sqrt{\cdot})'_{\mathbf{j}} d\delta_{\mathbf{j}}^2 - (h \circ \sqrt{\cdot})'_{\mathbf{k}} d\delta_{\mathbf{k}}^2 = 0 ;$$

given (2.4), this may be rewritten

$$(h \circ \sqrt{\cdot})'_{\mathbf{j}} (d\delta_{\mathbf{k}}^2 + 2 \delta_{\mathbf{j}\mathbf{k}} dx) - (h \circ \sqrt{\cdot})'_{\mathbf{k}} d\delta_{\mathbf{k}}^2 = 0 ,$$

from which

$$\left(\frac{\partial \delta_{\mathbf{k}}^2}{\partial x}\right)_{\Delta h} = 2 \delta_{\mathbf{j}\mathbf{k}} \frac{(h \circ \sqrt{\cdot})'_{\mathbf{j}}}{(h \circ \sqrt{\cdot})'_{\mathbf{k}} - (h \circ \sqrt{\cdot})'_{\mathbf{j}}} . \quad (3.10)$$

So we deduce from (3.9) and (3.10) that

$$\left(\frac{\partial y^2}{\partial x}\right)_{\Delta h} = 2 \left[\delta_{\mathbf{o}\mathbf{k}} \frac{(h \circ \sqrt{\cdot})'_{\mathbf{k}} + (h \circ \sqrt{\cdot})'_{\mathbf{j}}}{(h \circ \sqrt{\cdot})'_{\mathbf{k}} - (h \circ \sqrt{\cdot})'_{\mathbf{j}}} - x \right] . \quad (3.11)$$

The first consequence is that we are now able to compute the first of the derivatives we are looking for :

$$\left(\frac{\partial y}{\partial x}\right)_{\Delta h} = \frac{\delta_{\mathbf{o}\mathbf{k}} \frac{(h \circ \sqrt{\cdot})'_{\mathbf{k}} + (h \circ \sqrt{\cdot})'_{\mathbf{j}}}{(h \circ \sqrt{\cdot})'_{\mathbf{k}} - (h \circ \sqrt{\cdot})'_{\mathbf{j}}} - x}{y} \quad (3.12)$$

That formula immediately leads to the following properties (which may also be derived from Prop. 5.3):

(3.12) On $\mathcal{H}_+^2 - \{k\}$,

(a) When, on $]0, \infty[$, $h \circ \sqrt{\cdot}$ is derivable and convex,

$$(\partial y / \partial x)_{\Delta h} \leq 0 ;$$

(b) When, on $]0, \infty[$, $h \circ \sqrt{\cdot}$ is derivable and concave but h is convex,

$$(\partial y / \partial x)_{\Delta h} \geq 0 ;$$

(c) When, on $]0, \infty[$, $h \circ \sqrt{\cdot}$ is derivable and concave,

$$(\partial y / \partial x)_{\Delta h} \geq 0 \text{ if } x \leq \delta_{\mathbf{o}\mathbf{k}} .$$

The proof relies directly on (3.12), from which item (a) is obvious; when $h \circ \sqrt{\cdot}$ is concave, the sign of $(\partial y / \partial x)_{\Delta h}$ is the sign of

$(h \circ \sqrt{\cdot})'_k (\delta_{0k} - x) + (h \circ \sqrt{\cdot})'_j (\delta_{0k} + x)$, which expression yields item (c) and may be written [since $(h \circ \sqrt{\cdot})'(\delta^2) = h'(\delta) / 2\delta$] as

$$[h'_j \cos(\varphi_j) - h'_k \cos(\varphi_k)]/2,$$

yielding item (b) (see the proof of Prop. 3.3). **Q.E.D.**

Let us now tackle the second derivative. Considering again y^2 before y , we see from (3.8) that

$$\frac{\partial^2 y}{\partial x^2} = \frac{\frac{\partial^2 y^2}{2\partial x^2} - (\frac{\partial y}{\partial x})^2}{y}. \quad (3.13)$$

All we have yet to do is to find an expression of $(\partial^2 y^2 / \partial x^2)_{\Delta h}$. To simplify, we set $h = h \circ \sqrt{\cdot}$, $\sum h' = h'_j + h'_k$, and $\Delta h' = h'_j - h'_k$. We already know $(\partial \delta_k^2 / \partial x)_{\Delta h}$ from (3.10); a similar argument shows that

$$(\frac{\partial \delta_j^2}{\partial x})_{\Delta h} = -2 \delta_{jk} \frac{h'_k}{\Delta h'}.$$

We then compute from (3.11):

$$\begin{aligned} (\frac{\partial^2 y^2}{\partial x^2})_{\Delta h} &= 2 \left\{ -1 + \frac{\delta_{0k}}{\Delta h'^2} [\Delta h' (\bar{h}''_k \frac{2\delta_{jk} \bar{h}'_j}{\Delta h'} + \bar{h}''_j \frac{2\delta_{jk} \bar{h}'_k}{\Delta h'}) \right. \\ &\quad \left. + \sum \bar{h}' (\bar{h}''_k \frac{2\delta_{jk} \bar{h}'_j}{\Delta h'} - \bar{h}''_j \frac{2\delta_{jk} \bar{h}'_k}{\Delta h'}) \right\} \\ &= 2 \left\{ -1 + \frac{\delta_{jk}^2}{\Delta h'^3} [\bar{h}''_j \bar{h}'_k (\Delta h' - \sum \bar{h}') + \bar{h}''_k \bar{h}'_j (\Delta h' - \sum \bar{h}')] \right\} \\ &= 2 \left\{ -1 + \frac{2 \delta_{jk}^2}{\Delta h'^3} (\bar{h}''_k \bar{h}'_j^2 - \bar{h}''_j \bar{h}'_k^2) \right\} \\ &= 2 \left\{ -1 + \frac{2 \delta_{jk}^2 \bar{h}'_k^2 \bar{h}'_j^2}{\Delta h'^3} (\frac{\bar{h}''_k}{\bar{h}'_k^2} - \frac{\bar{h}''_j}{\bar{h}'_j^2}) \right\} \\ &= 2 \left(-1 + \frac{4 \delta_{jk}^2 \bar{h}'_k^2 \bar{h}'_j^2 \Delta g'}{\Delta h'^3} \right); \end{aligned} \quad (3.14)$$

.3.21.

in which we have noticed that $-h''/h'^2 = (1/h')' = 2[\sqrt{.}/(h'\sqrt{.})]'$, and defined $g = 1/2h' = \sqrt{.}/(h'\sqrt{.})$. We recognize here one of the functions mentioned in our commentary of Prop. 3.8 and 3.9. Last, we deduce from (3.12), (3.13), and (3.14) that

$$\left(\frac{\partial^2 y}{\partial x^2}\right)_{\Delta h} = \frac{1}{y} \left(\frac{4\delta_{ik}^2 \bar{h}'_k{}^2 \bar{h}'_j{}^2 \Delta g'}{\Delta \bar{h}'^3} - 1 \right) - \frac{1}{y^3} \left(\frac{\Sigma \bar{h}'}{\Delta \bar{h}'} \delta_{ok} + x \right)^2 . \quad (3.15)$$

4. Remarkable bounds and asymptotic behaviour: spatial extension

The market areas that appear for given values of Q and δ_{jk} and a given function h can be studied by simulation. We may draw the curve $Z_j \cap Z_k$ by means of a compass, when function h is easily invertible, using the formula $\delta_j = h^-(Q + h_k)$ for instance. We may also compute Δh at every point of a grid, then compare Δh to Q in order to see if the point belongs to Z_k or to Z_j , or both. We have implemented those methods on a computer to check our theoretical results and choose values of the parameters that should produce patterns of market areas in which the features we want to emphasize clearly appear.

Such an approach is insufficient, however, to predict the modification of the market areas induced by a change in function h . A purpose of the present work is precisely to show that apparently similar transportation cost functions may produce definitely different patterns of market areas. The analytical approach is then necessary to master the properties of our model, through the knowledge of the family of functions associated with each of them.

Another aspect of simulation is its double spatial limitation: in *extension* and *precision*. We cannot study the influence of the centres over an infinite area, neither can we decide if all the points of the area under study are more influenced by the one centre or the other. And the wider that investigation area, the smaller the density of the points considered. One may be satisfied with such a limited approach if it corresponds to the real spatial extension of the problem and to the spatial dispersion of the data. One may also wonder if a change in the scale of the study is likely to bring a change in the apparent properties. With both points of view in mind, we shall use the analytical approach in a way complementary to simulation; i.e., we shall study the properties of the market areas up to the infinitely great and up to the infinitesimal. That distinction *grosso modo* marks the separation between Chapters 4 and 5.

The present Chapter 4 is dedicated to the spatial extension of market areas. Essentially, the problem is of finding simple sets containing Z_k or contained in it, thereby constituting upper and lower bounds on Z_k . We begin by comparing it with the elementary sets $\emptyset$, $\{k\}$, and $\{j\}$, and by asking the

question whether Z_k is bounded; i.e., whether it is possible to encompass it with some circle. That is the matter of Sections 4.1, 4.2, and 4.3.

Thereafter in Sections 4.4, 4.5, and 4.7, we exploit the results of Chapter 3 in order to find circles, hyperbolæ, or vertical lines, delimitating Z_k more or less precisely. The aim is at becoming able to decide whether some characteristic regions are contained in Z_k or in Z_j , or contain Z_k , leaning only on the knowledge of some points or pieces of Z_k , Z_j , or $Z_j \cap Z_k$, and on some hypotheses about the transportation cost function. Of course the particular shapes of those bounds also yield some properties of the shape of the market areas themselves.

Hyperbolic and vertical bounds are particularly interesting when Z_k is not bounded. The Sections 4.6, 4.8 and 4.9 complete the information about this case. A first characteristic of the market areas is then the pair of symmetric directions taken by the demarcation line when the distance to the centres becomes infinite, and how those directions vary with Q and δ_{jk} . Then a question arises whether there are some points of the plane for which the acute angle formed by those directions contains Z_k or is contained in it; this leads us to the classical study of the possible straight asymptotes of the curve $Z_j \cap Z_k$, which also learns us if there is some distance above which $Z_j \cap Z_k$ is nearly linear. We have in fact to distinguish between a *limiting* line and an *asymptotic* one. The distance between the former and $Z_j \cap Z_k$ tends towards zero when the distance to the centers tends towards infinity. The latter is a limiting line which is not crossed any more by $Z_j \cap Z_k$ beyond some distance to the centres; an *asymptote* is a straight asymptotic line. This distinction between limiting and asymptotic lines is only used to express our lack of knowledge regarding some cases. Our investigation also allows us to know if the boundaries of the hyperbolic and vertical bounds are asymptotic to $Z_j \cap Z_k$, when those bounds may thus be considered as tight.

4.1. Emptiness of market area Z_k

(4.1) (a) $Z_k \neq \emptyset \Leftrightarrow Z_k \cap [\delta_{ok}, \infty[\times \{0\} \neq \emptyset$;

(b) If h is concave, then: $Z_k \neq \emptyset \Leftrightarrow k \in Z_k$;

(c) If h is convex, then: $Z_k \neq \emptyset \Leftrightarrow Z_k$ is unbounded.

The proposition still holds if ' Z_k ' is replaced everywhere by ' $Z_k - Z_j$ '.

We give the proof concerning Z_k only; the one regarding $Z_k - Z_j$ is similar. The second implication ($\Leftarrow$) of item (a) is quite plain. The first ($\Rightarrow$) relies on Prop. 3.1: if Z_k contains some point i , then it also contains the intersection point i' of the circle passing through i and centred on k with that part of the x -axis which lies at the right of k ; i.e., $[\delta_{ok}, \infty[\times \{0\}$. Accordingly, the study about the emptiness of Z_k may be restricted to that set. As we have there $\delta_j = x + \delta_{ok}$ and $\delta_k = x - \delta_{ok}$, the derivative of Δh w.r.t. x is:

$$(\partial \Delta h / \partial x)_{y=0} = h'_j - h'_k ,$$

which expression is negative when h is concave and positive when h is convex. In the former case, Δh is decreasing w.r.t. x . Consequently if point i lies in Z_k at the right of k on the x -axis, i.e., if $i \in Z_k \cap [\delta_{ok}, \infty[\times \{0\}$, then $[ki] \subseteq Z_k$ since ' $\Delta h \geq Q$ ' defines Z_k ; so a fortiori $k \in Z_k$. In the latter case, Δh is increasing w.r.t. x , so that ' $i \in Z_k \cap [\delta_{ok}, \infty[\times \{0\}$ ' entails that $[x, \infty[\times \{0\} \subseteq Z_k$, and Z_k is unbounded. The implications ' $\Leftarrow$ ' of items (b) and (c) being obvious, the proof is complete. Q.E.D.

Whether the conditions ' $k \in Z_k$ ' or ' Z_k is unbounded' are met is easy to check. The former is equivalent to ' $\Delta h_k \geq Q$ '. This means in particular that Z_k is never empty when $h = \ln$, for we have then $\Delta h_k = \infty$; when $h(\delta) = \delta$, on the contrary, Z_k vanishes when $\delta_{jk} < Q$. Let us now study the latter condition.

4.2. Boundedness of market area Z_k

A set E is *bounded* iff. its diameter is finite; in other words, iff. it is contained in some disk. Market area Z_j is of course unbounded, as $\mathbb{R}_+ \times \mathbb{R} \subset Z_j$. As to the boundedness conditions on Z_k , they are related to the limiting value of Δh when point i moves towards infinity, since ' $\Delta h \geq Q$ ' defines Z_k . This is the philosophy of the next Prop. 4.2. That result emphasizes the importance of the ratio δ_{jk}/Q (this was confirmed by Hanjoul et al., Section 10.2, by the study of an approximation of market areas when $h(\delta) = \delta^a$).

From now on we shall denote by $h'(\infty)$ the limiting value $\lim_{\delta \rightarrow \infty} h'(\delta)$, if any.

(4.2) (a) Z_k is unbounded $\Leftrightarrow Z_k \cap [\delta_{ok}, \infty[\times \{0\}$ is unbounded;

(b) If $\delta_{jk} h'(\infty) > Q$, then $Z_k - Z_j$ is unbounded;

$< Q$, then Z_k is bounded;

(c) If $\delta_{jk} h'(\infty) = Q$ and h is concave (resp. strictly concave),
 Z_k (resp. $Z_k - Z_j$) is unbounded;

(d) If $\delta_{jk} h'(\infty) = Q$ and h is convex, then the next equivalences hold for some $\delta^* \geq 0$:

Z_k is unbounded

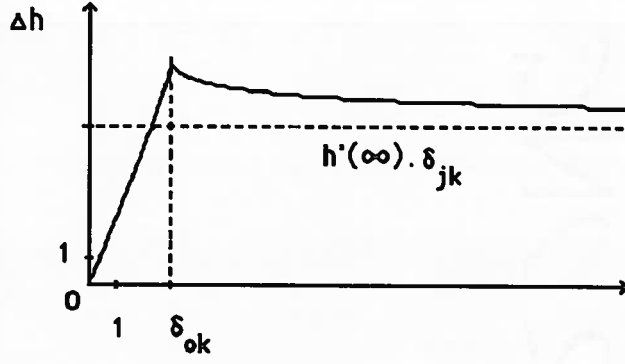
$\Leftrightarrow Z_k \neq \emptyset$

$\Leftrightarrow h$ is linear on $[\delta^*, \infty[$ and nonlinear on $[\delta^{**}, \infty[$ for any $\delta^{**} \in]0, \delta^*[$ ¹

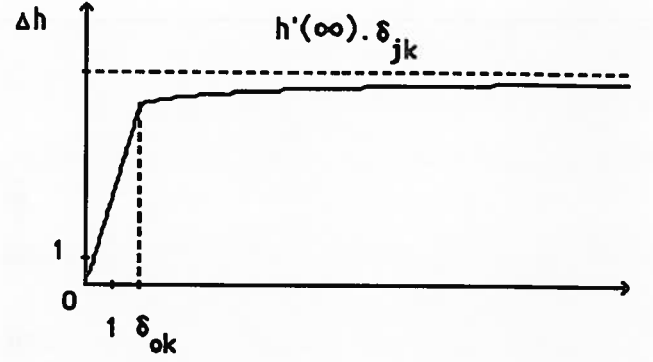
$\Leftrightarrow Z_k$ is the straight half-line $[\delta_{ok} + \delta^*, \infty[\times \{0\}$;

(e) If $\delta_{jk} h'(\infty) = Q$ and h is strictly convex, Z_k is empty and thus bounded.

¹ 'Nonlinear' means of course here 'not (globally) linear'; h may be linear on subintervals of $[\delta^{**}, \infty[$.



(a) concave h : $h(\delta) = \delta + \sqrt{\delta}$, $\delta_{0k} = 3$



(b) convex h : $h(\delta) = 2\delta - \sqrt{1 + \delta}$, $\delta_{0k} = 2$

Fig. 4.1. Evolution of the transportation cost difference with x along the x -axis.

The proof of item (a) is similar to that of Prop. 4.1a. As to item (b) (Fig. 4.1), the limiting value of Δh when point i moves towards infinity along the x -axis in the positive direction is given by $\delta_{jk} h'(\infty)$. This is obvious from the mean-value theorem, which states that for any i on the half-axis $[\delta_{0k}, \infty[\times \{0\}$ we have

$$\Delta h = h'(\delta) (\delta_j - \delta_k) = h'(\delta) \delta_{jk}$$

for some $\delta \in]\delta_k, \delta_j[$. Accordingly, when $\delta_{jk} h'(\infty) > Q$, there is some $\delta^- \geq 0$ for which we must have $\Delta h = h(\delta_k + \delta_{jk}) - h(\delta_k) > Q$, $\forall \delta_k \geq \delta^-$, as h and Δh are continuous functions. The straight half-line $[\delta_{0k} + \delta^-, \infty[\times \{0\}$ is consequently included in $Z_k - Z_j$, which is thus unbounded. The proof of the boundedness of Z_k when $\delta_{jk} h'(\infty) < Q$ is similar. As to point (c), function h being concave, $h'(\delta) \geq h'(\infty)$; so, for any $i \in [\delta_{0k}, \infty[\times \{0\}$, we have

$$\Delta h = h'(\delta) \delta_{jk} \geq h'(\infty) \delta_{jk} = Q.$$

The whole straight half-line $[\delta_{0k}, \infty[\times \{0\}$ is then a part of Z_k . When function h is strictly concave, we have $\Delta h > Q$ on the set $[\delta_{0k}, \infty[\times \{0\}$, which is consequently included in $Z_k - Z_j$. Item (e) directly derives from item (d).

We now come to the chain of equivalences listed in (d). The first one, ' Z_k is unbounded $\Leftrightarrow Z_k \neq \emptyset$ ', has already been proved: see Prop. 4.1c. Here follow the other, in which ' Z_k is unbounded' is accordingly replaced by ' $Z_k \neq \emptyset$ ':

i) $Z_k \neq \emptyset \Leftrightarrow h$ is linear beyond some δ^* . (Due to the continuity of h , the assertions ' $\exists \delta^*: h$ is linear on $[\delta^*, \infty[$ ' and ' $\exists \delta^*: h$ is linear on $[\delta^*, \infty[$ and nonlinear on $[\delta^{**}, \infty[\forall \delta^{**} \in]0, \delta^*[$ ' are equivalent). ($\Rightarrow$) As $Z_k \neq \emptyset$, some point i' on $[\delta_{ok}, \infty[\times \{0\}$ belongs to Z_k (Prop. 4.1a); so $\Delta h_{i'} \geq Q$. As Δh is an increasing function of x on $[\delta_{ok}, \infty[\times \{0\}$ (see Prop. 3.3c), we have also $\Delta h_{i'} \leq h'(\infty) \delta_{jk} = Q$. Consequently $\Delta h_{i'} = \Delta h_i = Q$ for every $i \in [x_{i'}, \infty[\times \{0\}$; i.e.,

$$h(\delta_{jk} + \delta) - h(\delta) = Q \quad \forall \delta \geq \delta_{ki}.$$

As h is convex, this implies the linearity of h for $\delta \geq \delta_{ki}$. ($\Leftarrow$) If h is linear for $\delta \geq \delta^*$, Δh is constant on $[\delta_{ok} + \delta^*, \infty[\times \{0\}$ and its value there is $\delta_{jk} h'(\infty) = Q$. That straight half-line is then necessarily in Z_k , and $Z_k \neq \emptyset$.

ii) h is linear for $\delta \geq \delta^*$ and nonlinear on $[\delta^{**}, \infty[\forall \delta^{**} \in]0, \delta^*[\Leftrightarrow Z_k = [\delta_{ok} + \delta^*, \infty[\times \{0\}$. ($\Rightarrow$) The linearity of h for $\delta \geq \delta^*$ implies the constancy of Δh on $[\delta_{ok} + \delta^*, \infty[\times \{0\}$, and its value there is $\delta_{jk} h'(\infty) = Q$; so that this set is in Z_k (in $Z_k \cap Z_j$, more precisely). Furthermore, all the points that do not belong to that half-line lie outside Z_k . Consider first a point i of the x -axis between k and $(\delta_{ok} + \delta^*, 0)$, the origin of the half-line; i.e., $i \in [k, (\delta_{ok} + \delta^*, 0)[$. Then the convexity of h implies

$$\Delta h = h(\delta_{jk} + \delta_k) - h(\delta_k) < h(\delta_{jk} + \delta^*) - h(\delta^*) = h'(\infty) \delta_{jk} = Q,$$

the inequality being strict because of the nonlinearity of h on $[\delta_k, \infty[$. So $i \notin Z_k$. Second, take $i \notin [\delta_{ok}, \infty[\times \{0\}$. According to the proof of Prop. 4.1a, $\Delta h < \Delta h_{i'}$ for some $i' \in [\delta_{ok}, \infty[\times \{0\}$; as we have just seen that $\Delta h_{i'} \leq Q$, we have $\Delta h < Q$, and $i \notin Z_k$. ($\Leftarrow$) As Δh is continuous, the fact that Z_k is a straight half-line entails that $Z_k = Z_j \cap Z_k$. That is to say, $\Delta h = Q$ on Z_k ; or $h(\delta_{jk} + \delta) - h(\delta) = Q \quad \forall \delta \geq \delta^*$. Being a convex function, h must then be linear on $[\delta^*, \infty[$. And if it were also linear on $[\delta^{**}, \infty[$ with $0 < \delta^{**} < \delta^*$, Δh should be constant on $[\delta_{ok} + \delta^{**}, \infty[$, which should then be a part of Z_k ; whereas it is not. This completes the proof. **Q.E.D.**

Let us see how those statements apply when $h(\delta) = \delta^a \quad \forall \delta \geq 0$, with $a > 0$ (Fig. 3.4). If $0 < a < 1$, we have $h'(\infty) = 0$; according to Prop. 4.2b, Z_k is then bounded. If $a > 1$, $h'(\infty) = \infty$: Z_k is unbounded (notice however that the shape of Z_k is not the same when $1 < a < 2$ as when $a > 2$). Last, when $a = 1$, $h'(\infty) = 1$. In this case, if $\delta_{jk} > Q$, Z_k is unbounded; while $Z_k = \emptyset$ if $\delta_{jk} < Q$

(see Prop. 4.1c). If $\delta_{jk} = Q$, we may refer to Prop. 4.2c as well as to Prop. 4.2d, as h is now linear. Both indicate that Z_k is unbounded, but Prop. 4.2d tells us, moreover, that Z_k degenerates into the straight half-line $[\delta_{ok}, \infty[\times \{0\}$.

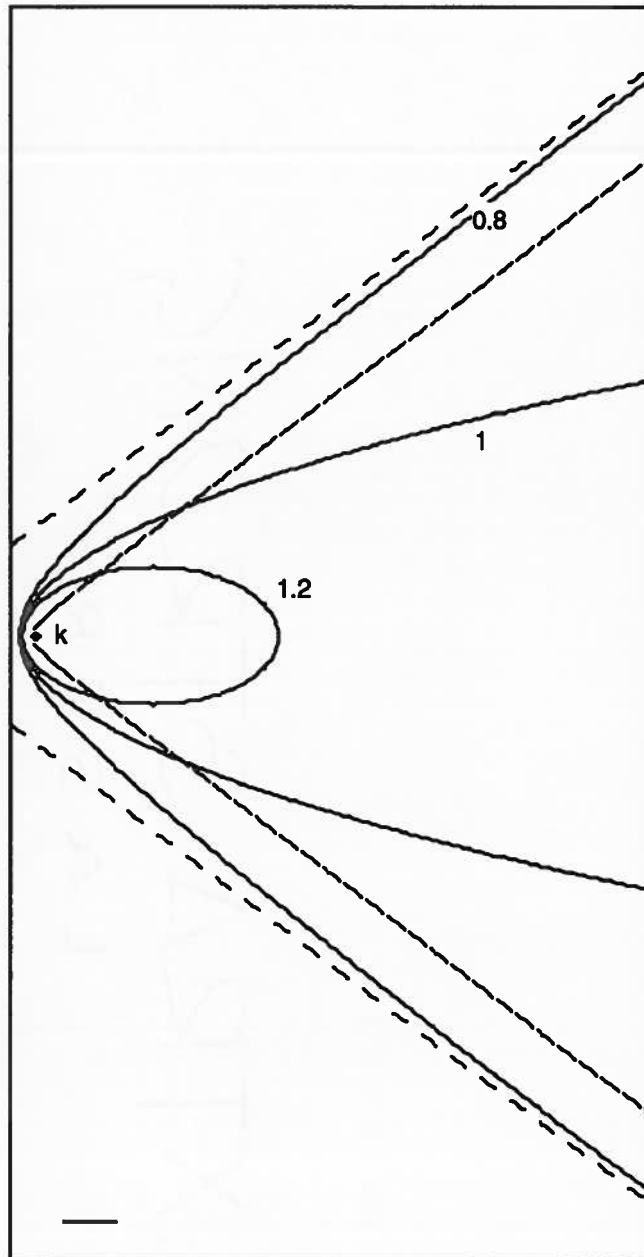


Fig. 4.2. — Market areas generated by the t.c.f. $\delta + \ln(\delta)$ when $\delta_{jk} = 1$ for indicated values of Q .

When $Q = 1$, the indifference line has no straight limiting line.

When $Q < 1$, it has two oblique ones crossing at the point $(-Q/\sqrt{1-Q^2}, 0)$.

An example of a more complex case is given by the transportation cost function $h(\delta) = \delta + \ln(\delta)$; see Fig. 4.2. Here, market area Z_k is bounded (and nonempty) when $Q > \delta_{jk}$ and unbounded when $Q \leq \delta_{jk}$. That t.c.f. corresponds to the distance deterrence function $\exp(-\beta\delta)/\delta^2$, a particular case of the function $\exp(-\beta\delta)/\delta^{\beta'}$, intermediate between The Pareto and negative exponential functions and used by Wardrop (1959), Tanner (1961) and Harris (1964).

4.3. Position of centres relatively to market areas

This section is related with the two preceding ones, 4.1 and 4.2, as well as with Sections 5.1 and 5.2, in that the field of study can be reduced to the x-axis; more precisely, as appears from the next proposition, to the half-line $[\delta_{ok}, \infty[\times \{0\}$. As to centre j , we already know from Prop. 2.3 that it belongs to its own market area; this is not necessarily true for centre k : see Fig. 3.4.

(4.3) (a) Centre $k \in Z_k$ iff. $\ell \in]ok]$;

(b) If h is concave and $Z_k \neq \emptyset$, then $k \in Z_k$;

(c) When h is convex and nonlinear, then $]h(\delta_{jk}) - h(0), h'(\infty)\delta_{jk}[$ is not empty and, if Q belongs to that interval, $k \notin Z_k \neq \emptyset$.

According to Prop. 3.3a, we have indeed $\Delta h < \Delta h_k$ for any $i \in]ok[$. Consequently if $k \notin Z_k$, i.e., if $\Delta h_k < Q$, then $\Delta h < Q$ and $]ok] \cap Z_k = \emptyset$: so $\ell \notin]ok]$. If $k \in Z_k$, $]ok] \cap Z_k \neq \emptyset$; and as $\ell \in \mathcal{R}_+^* \times \{0\}$ (Prop. 2.3), we see that $\ell \in]ok]$. Item (a) is thus proved. Item (b) repeats Prop. 4.1b. As to item (c), it is obvious that $k \notin Z_k$ iff. $\Delta h_k < Q$, and that $Z_k \neq \emptyset$ if $Q < \delta_{jk} h'(\infty)$ (Prop. 4.1c and 4.2b). The only problem is to prove that $\Delta h_k < \delta_{jk} h'(\infty)$. As Δh is increasing w.r.t. x on $[\delta_{ok}, \infty[\times \{0\}$ because h is convex (Prop. 3.3b), and as we have seen that $\delta_{jk} h'(\infty)$ is the limiting value of Δh on the right of the x-axis, we should have $\Delta h_k = \delta_{jk} h'(\infty)$ only if Δh were constant on $[\delta_{ok}, \infty[\times \{0\}$. But this would mean that $h(\delta + \delta_{jk}) - h(\delta) = Q \forall \delta \geq 0$; as h is convex, this would entail the linearity of that same function. Item (c) is so verified *ab absurdo*. **Q.E.D.**

² According to March (1971), that function was used by Johnsson (1952). March does not explain clearly, unfortunately, whether Johnsson's model was a GG model or not. That kind of ambiguity pervades whole March's paper.

4.4. Circular bounds

The main result is here the one that follows from Prop. 3.8. For any $i \in Z_j \cap Z_k$, we call D_i the disk equal to G_i or to $\mathbb{R}^2 - G_i^\circ$ in Section 3.5 when $h = (.)^2$, if $\alpha_i \neq 1$. When i is not a point of the x -axis, D_i is the disk tangent to $Z_j \cap Z_k$ at i and the centre of which lies on the x -axis. On the other hand, when Z_k is bounded, $D[(\ell + r)/2, \ell]$, which is equal to $D[(\ell + r)/2, r]$, is the disk of which $[\ell r]$ is a diameter.

(4.4) (a) When $\ln[h^\wedge/(h' \cdot h^\wedge)]$ is concave on $h > 0$, $\infty[>$,

(i) if Z_k is bounded, $Z_k \subseteq D[(\ell + r)/2, r]$;

(ii) if $\ln[h^\wedge/(h' \cdot h^\wedge)]$ is strictly increasing on $h > 0$, $\infty[>$,

$D_i \subseteq Z_k \forall i \in Z_j \cap Z_k$, and $Z_k = \cup \{D_i; i \in Z_j \cap Z_k\}$;

(iii) $\ln[h^\wedge/(h' \cdot h^\wedge)]$ is not strictly decreasing on $h > 0$, $\infty[>$;

(b) When $\ln[h^\wedge/(h' \cdot h^\wedge)]$ is convex on $h > 0$, $\infty[>$,

(i) if Z_k is bounded, $D[(\ell + r)/2, r] \subseteq Z_k$;

(ii) if $\ln[h^\wedge/(h' \cdot h^\wedge)]$ is strictly increasing on $h > 0$, $\infty[>$,

$Z_k \subseteq D_i \forall i \in Z_j \cap Z_k$, and $Z_k = \cap \{D_i; i \in Z_j \cap Z_k\}$;

(iii) if $\ln[h^\wedge/(h' \cdot h^\wedge)]$ is strictly decreasing on $h > 0$, $\infty[>$,

$D_i \subseteq Z_j, \forall i \in Z_j \cap Z_k$, and $Z_j = \cup \{D_i; i \in Z_j \cap Z_k\}$.

Prop. 4.4 is a mere application of Prop. 3.8. The only problem is to know whether G_i and $G_{\ell r}$, the auxiliary t.c.f. $h^\sim$ being equal to $(.)^2$, correspond to D_i and $D[(Q + r)/2, r]$ or to the complementary areas of their interiors. As to $G_{\ell r}$, we have $G_{\ell r} = D[(\ell + r)/2, r]$ in any case.

We have now to see whether G_i is equal to D_i or to the closure of its exterior. According to inequality (3.7), the problem boils down to knowing whether $\alpha_i < 1$ or $\alpha_i > 1$. Equality (3.5) shows that $\alpha_i < 1$ when $h^\sim'(\delta)/h'(\delta)$ is strictly increasing w.r.t. δ^2 (or δ), and that $\alpha_i > 1$ when it is strictly decreasing. The proof of Prop. 4.4 is then immediate. **Q.E.D.**

In particular, the case where $\delta/h'(\delta)$ would be concave and strictly decreasing w.r.t. δ^2 is impossible, for it would lead to: $Z_j \subseteq D_i$. However, $\delta/h'(\delta)$ may have that property when δ is in an open interval containing $[\delta_k, \delta_j]$. The inclusion ' $Z_j \subseteq D_i$ ' then makes sense in a vicinity of point i : there is some vicinity V of i (i.e., a set including some open set containing i) for which $Z_j \cap V \subseteq D_i \cap V$. This is what happens in Fig. 5.1b if $\delta_k \leq 3\delta^*/5$. Similar remarks can be made for all the items of Prop. 4.4. Also, relatively to any particular point i of $Z_j \cap Z_k$, the condition that $\delta/h'(\delta)$ should be strictly increasing or decreasing may be replaced by the local one that $\alpha_i < 1$ or $\alpha_i > 1$. A final note: when $\alpha_i = 1$, the area $\equiv \delta_k^2 \leq \alpha_i \delta_j^2 + \zeta_i$ degenerates into the half-plane $[x, \infty[\times \{0\}$; it should not be difficult for the reader to foresee what ensues when $\delta/h'(\delta)$ is concave or convex w.r.t. δ^2 .

Prop. 4.4 has the interesting consequence that we are able to get information about the smallest disk circumscribed to Z_k and about a largest disk inscribed in Z_k . Let us denote by s (a *summit*) some point maximizing the ordinate y on Z_k , if any; and by s' its orthogonal projection on the x -axis.

(4.5) (a) If $\ln^\circ[h^\wedge/(h' \cdot h^\wedge)]$ is concave on $h > 0, \infty[>$,

- (i) if Z_k is bounded, $D[(\ell+r)/2, r]$ is the smallest disk containing Z_k ;
- (ii) if a summit s exists and $\ln^\circ[h^\wedge/(h' \cdot h^\wedge)]$ is strictly increasing on $h > 0, \infty[>$, $D(s', s)$ is a largest disk contained in Z_k ;

(b) If $\ln^\circ[h^\wedge/(h' \cdot h^\wedge)]$ is convex on $h > 0, \infty[>$,

- (i) if Z_k is bounded, $D[(\ell+r)/2, r]$ is a largest disk contained in Z_k ;
- (ii) if $\ln^\circ[h^\wedge/(h' \cdot h^\wedge)]$ is strictly increasing on $h > 0, \infty[>$, a unique summit s exists and $D(s', s)$ is the smallest disk containing Z_k .

(The proof, being obvious, is omitted). The bounds recorded in Prop. 4.4 are *tight* in the sense that, for Prop. 4.4a, for instance, D_i is the largest disk contained in Z_k and the circle of which passes through i ; etc. This is

guaranteed by the symmetry of Z_k w.r.t. the x -axis. Let us detail the case of item (a) for instance. If point i is not on the x -axis, a disk contained in Z_k , tangent to $Z_j \cap Z_k$ at i , but larger than D_i , would have its centre lying in Z_k on the normal to $Z_j \cap Z_k$ at i (ie., on the straight line joining i to the centre of D_i), but farther from i than the centre of D_i ; and so it would encompass the symmetric of i w.r.t. to the x -axis, and contain points of $Z_j - Z_k$: a contradiction.

For continuity reasons, this remains true when i is on the x -axis. More generally, anyway, it is obvious that

(4.6) *If a point i lies on the x -axis, D_i is the osculating disk to $Z_j \cap Z_k$ at i .*

I.e., the disk the circle of which is locally most alike to $Z_j \cap Z_k$ at i ; the radius of D_i is then the radius of curvature of $Z_j \cap Z_k$ at i . This may be useful when drawing $Z_j \cap Z_k$.

We now give two other circular bounds on Z_k . They are necessarily less precise than the above-mentioned ones, but the conditions for their application are different from those of Prop. 4.4. We denote by A_p the Apollonian disk relative to points j and k and the circle of which passes through a point p ; the area A_p is described by the inequality $\delta_j/\delta_k \geq \delta_{jp}/\delta_{kp}$. When $p \in \mathfrak{X}_+ * \mathfrak{X}$, A_p is a disk; it is the closure of the exterior of a disk when $p \in \mathfrak{X}_- * \mathfrak{X}$. The abscissa of the centre κ_i of a disk A_i is

$$x_{\kappa_i} = \delta_{ok} \frac{\delta_j^2 + \delta_k^2}{\delta_j^2 - \delta_k^2} = \frac{\delta_o^2 + \delta_{ok}^2}{2x}. \quad (4.1)$$

When Z_k is empty, we set $D(k, r) = D(k, \ell) = A_r = A_\ell = \emptyset$.

(4.7) (a) *If Z_k is bounded, $Z_k \subseteq D(k, r)$;*

(b) *If $k \in Z_k$, $D(k, \ell) \subseteq Z_k$;*

(c) *$D(j, \ell) \subset Z_j$ ³.*

³ For concave or convex t.c.f.'s, Prop. 4.7 appears as a particular case of a more general property: above the x -axis, if h is concave, δ_k decreases when φ_k increases; if h is convex, δ_j increases when φ_j increases. We leave the proof to the reader.

(4.8) (a) If $h \circ \exp$ is convex, then

(i) if Z_k is bounded, $Z_k \subseteq A_r$;

(ii) if $k \in Z_k$, $A_\ell \subseteq Z_k$;

(b) If $h \circ \exp$ is concave, then $Z_k \subseteq A_\ell$ and, if Z_k is bounded, $A_r \subseteq Z_k$.

Those properties obviously derive from Prop. 3.1 and 3.5a, and from the fact that, when Z_k is not empty, centre k either belongs to Z_k or to $]o\ell[$ (Prop. 4.3a). Also, regarding Prop. 4.8b, we know from Prop. 3.7a that the concavity of $h \circ \exp$ entails the concavity of h itself; which in turn implies that $k \in Z_k$ if $Z_k \neq \emptyset$ (Prop. 4.3b). Here is for instance a proof of Prop. 4.7c. For every $p \in [j\ell]$, the circle centred on j and passing through ℓ is in Z_j according to Prop. 3.1. As $D(j, \ell)$ is the union of all such circles, we have finally $D(j, \ell) \subset Z_j$. Q.E.D.

Prop. 4.4 and 4.5 show that the maximal distance between any two points of Z_k , i.e., its diameter Φ_k , may be equal to the length L_x of the projection of Z_k upon the x-axis or to the length L_y of its projection upon the y-axis. From another point of view, the ratio L_x/L_y may be < 1 or > 1 , which means that Z_k is somehow directed as the segment $[jk]$ or perpendicular to it. When $k \in Z_k$, Prop. 4.7b and 4.8 may be used to produce additional information about the ratios L_x/L_y and Φ_k/L_x . Notice that L_x and L_y are $\leq \Phi_k$, so that $\Phi_k/L_x \geq \max \{1, L_y/L_x\}$.

COROLLARY 4.1 When $L_x \notin \{0, \infty\}$,

(a) if $\ln \circ [h^\wedge / (h' \circ h^\wedge)]$ is concave and strictly increasing,

then $L_x = \Phi_k$, and $L_y/L_x \leq 1$ when L_y is finite;

(b) if $\ln \circ [h^\wedge / (h' \circ h^\wedge)]$ is convex and strictly increasing,

$L_y = \Phi_k$ and $1 \leq L_y/L_x$;

(c) if $k \in Z_k$, $L_y/L_x \leq \Phi_k/L_x < 2$;

if, moreover, $h \circ \exp$ is convex,

$L_y/L_x \leq \Phi_k/L_x \leq 1 + \delta_{ok}/(\delta_{ok} + L_x)$;

(d) if $h \circ \exp$ is concave and $L_x \geq \delta_{ok}$,

$1 - \delta_{ok}^2/L_x^2 \leq L_y/L_x \leq \Phi_k/L_x < 2$.

Items (a) and (b) derive from Prop. 4.5. As to item (c), if $k \in Z_k$, we have $\delta_{kr} \leq L_x$; and the fact that $L_x \in \mathcal{R}_+^*$ implies that Z_k is bounded and that $Z_k \subset D(k, r)$, so that $\Phi_k < 2 \delta_{kr}$ (the inclusion is strict as $k \neq r$: see Prop. 3.1; and k is $\neq r$ because ' $k = r$ ' would imply that $Z_k = D(k, k) = \{k\}$ and that $L_x = 0$). When $h \circ \exp$ is convex, $Z_k \subseteq A_r$, so that Φ_k is not larger than the diameter of A_r , which is $x - \delta_{ok}^2/x_r$. As that expression is increasing w.r.t. x_r and as ' $k \in Z_k$ ' implies that $x_r \leq \delta_{ok} + L_x$, the proof of (c) is completed. When $h \circ \exp$ is concave, $A_r \subseteq Z_k$ and $L_y \geq x_r - \delta_{ok}^2/x_r$. The hypothesis that $L_x \geq \delta_{ok}$ implies that the smallest possible value of x_r is L_x . Hence the lower bound on L_y/L_x . On the other hand, h is now concave (Prop. 3.7a) and Z_k is not empty since $L_x > 0$; so item (c) yields the upper bound on Φ_k/L_x (the specification ' $L_x < \infty$ ' is useless because it follows from the concavity of $h \circ \exp$ that $h'(\infty) = 0$, from which we know that Z_k is bounded: see Prop. 4.2b). **Q.E.D.**

A second consequence of Prop. 4.7 concerns the points ℓ and r (see also Prop. 2.4):

COROLLARY 4.2. *Point ℓ is the point of $Z_j \cap Z_k$ closest to $[jk]$; if Z_k is bounded, point r is the point of $Z_j \cap Z_k$ farthest from $[jk]$.*

Let us now have a look at two examples of application of all those properties. Take first the monomial transportation cost function: δ^a , with $a > 0$. Here $\delta/h'(\delta)$ is equal to δ^{2-a}/a , or $(\delta^2)^{1-a/2}/a$, so that, for every $i \in Z_j \cap Z_k$, $D_i \subseteq Z_k$ if $a \in]0, 2]$ and $D_i \subset Z_j$ if $a > 2$. The abscissa of the centre c_i of D_i is

$$x_{c_i} = \delta_{ok} \frac{\delta_j^{2-a} + \delta_k^{2-a}}{\delta_j^{2-a} - \delta_k^{2-a}}. \quad (4.2)$$

The similarity between (4.1) and (4.2) is not casual, as $\ln(\delta)$ is a kind of limit of δ^a when $a \rightarrow 0$; see Hanjoul et al., Chapter 10. On the other hand, $h(\delta) = \delta^a = \exp[a \ln(\delta)]$ is convex w.r.t. $\ln(\delta)$; if $a < 1$, in particular, we thus have $A_\ell \subseteq Z_k \subseteq A_r$.

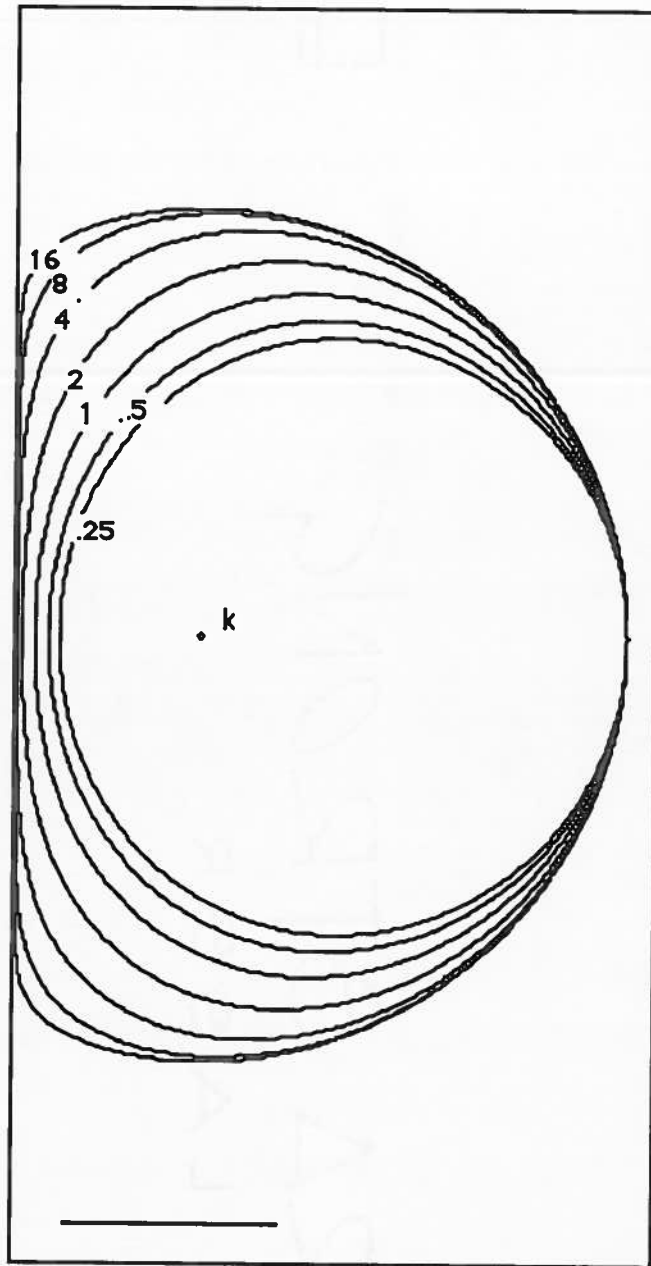


Fig. 4.3a. — Market areas generated by the t.c.f. $-\delta^{-a}$ when the indifference line is constrained to pass through the point $(2.875, 0)$; $\delta_{jk} = 1.75$, the exponent a is as indicated.

A second example is $h(\delta) = -\delta^{-a}$, with $a > 0$; see Fig. 4.3a. In opposition with the previous one, $\delta/h'(\delta)$ is now $(\delta^2)^{1+a/2}/a$, so that, for every $i \in Z_j \cap Z_k$, we have $Z_k \subseteq D_i$. The ratio L_y/L_x is consequently now in the range $]1, 2[$, instead of $]0, 1[$; notice that it is easy to prove that Z_k tends towards $A_p - \mathcal{R}^* \times \mathcal{R}$ when $a \rightarrow \infty$ and $Z_j \cap Z_k$ is constrained to pass through a

fixed point p . As to $h \circ \exp$, it is now concave, as $h(\delta) = -\exp[a \ln(\delta)]$; so that $A_r \subseteq Z_k \subseteq A_\theta$.

Other examples illustrating the possible values of L_y/L_x are given in Fig. 4.3b and 4.3c.

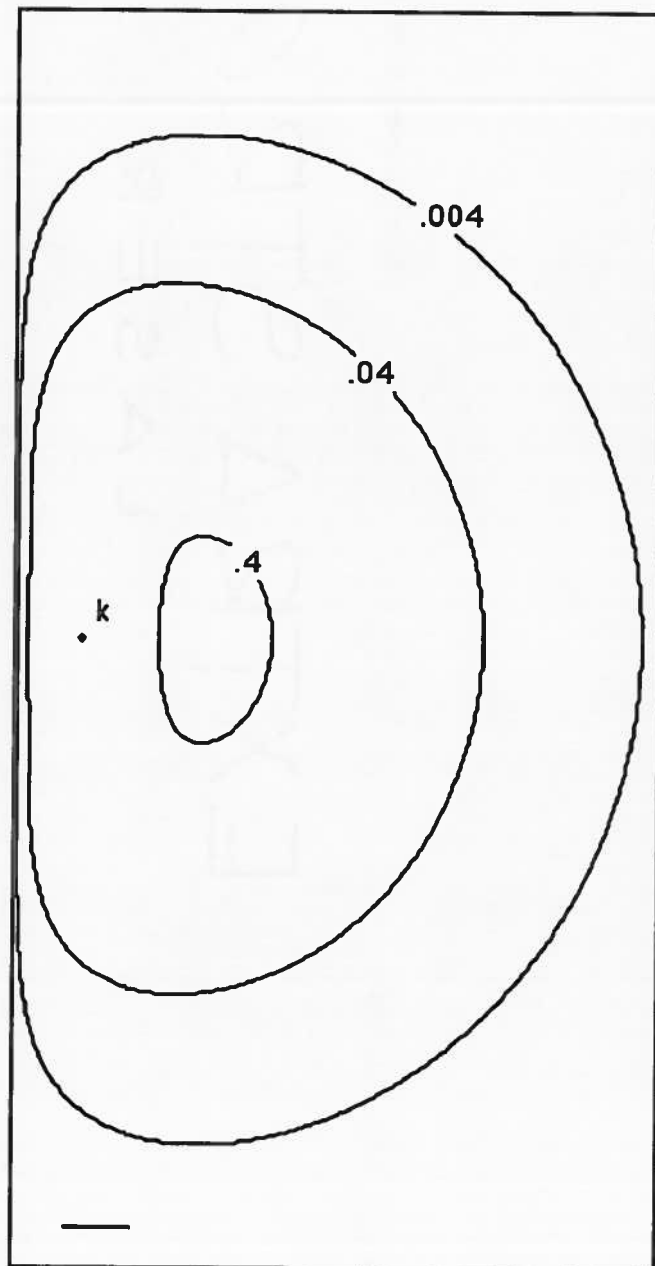


Fig. 4.3b — Market areas generated by the t.c.f. $1/[1 + \exp(3 - \delta)]$; $\delta_{jk} = 2$, Q as indicated.

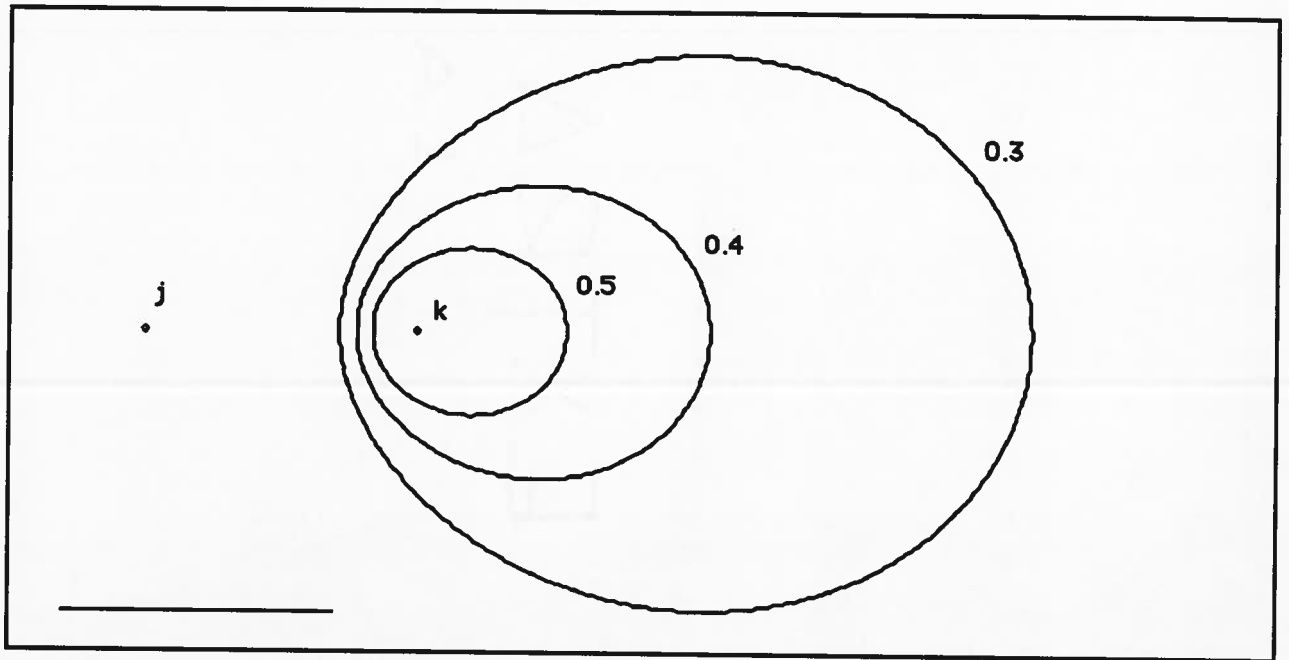


Fig. 4.3c. — Market areas generated by the t.c.f. $\sqrt{\delta}$; $\delta_{jk} = 1$, Q as indicated. This appears in the literature as the 'square root' or 'leptokurtic' exponential case.

4.5. Hyperbolic bounds

We define the two following hyperbolic areas:

$$H_{\ell} = \{i: \delta_j - \delta_k \geq \delta_{j\ell} - \delta_{k\ell} \text{ and } \delta_j \geq \delta_{j\ell}\}$$

$$H_{\infty} = \{i: \delta_j - \delta_k \geq Q/h'(\infty)\}.$$

To reach a full understanding of the proposition that follows, it is necessary to notice that those areas may in some cases degenerate into obvious bounds, making the proposition useless. To simplify the formulation of the next proposition, if $h'(\infty) \notin \mathbb{R}_+ \cup \{\infty\}$, we set $H_{\infty} = \emptyset$ if h is concave, and $H_{\infty} = \mathbb{R}^2$ when h is convex.

When $\ell \in]ok]$, the boundary of H_{ℓ} is a branch of a hyperbola (degenerating into $[\delta_{ok}, \infty[\times \{0\}$ if $k = \ell$) with foci j and k ; the specification that $\delta_j \geq \delta_{j\ell}$ in the definition of H_{ℓ} is superfluous because ℓ is then the closest point of the branch to centre j (Corollary 4.2). But it may happen (Prop. 4.3a and c) that $k \in]o\ell[$; in that case H_{ℓ} degenerates into $[x_{\ell}, \infty[\times \{0\}$, and the specification that $\delta_j \geq \delta_{j\ell}$ is now necessary. The information yielded

by Prop. 4.9b about H_ℓ when those cases of degeneracy occur is nothing but a direct implication of Prop. 3.3b and c.

The definition of H_∞ shows that it degenerates into $\mathbb{R}_+ \times \mathbb{R}$ if $h'(\infty) = \infty$. It obviously also degenerates into the straight half-line $[\delta_{ok}, \infty[\times \{0\}$ when $Q = h'(\infty) \delta_{jk}$, and becomes empty when $Q > h'(\infty) \delta_{jk}$.

(4.9) (a) If h is concave, $H_\infty \subseteq Z_k \subseteq H_\ell$;

(b) If h is convex, $H_\ell \subseteq Z_k \subseteq H_\infty$.

The proof concerning H_ℓ directly derives from Prop. 3.5b. As to H_∞ , the mean-value theorem tells us that $h_j - h_k = h'(\delta) (\delta_j - \delta_k)$ for some $\delta \in]\delta_k, \delta_j[$. If h is concave, for instance, we have $h'(\delta) \geq h'(\infty)$, so that $h_j - h_k \geq h'(\infty) (\delta_j - \delta_k)$: hence, the fact that $h'(\infty) (\delta_j - \delta_k) \geq Q$ implies that $h_j - h_k \geq Q$; i.e., $H_\infty \subseteq Z_k$. (The argument is similar when h is convex). Q.E.D.

Let us study how Prop. 4.9 can be used for the monomial transportation cost function. The bound H_∞ is here completely useless: when $a \in]0, 1[$, it degenerates into $\emptyset$; when $a = 1$, it is equal to Z_k ; and when $a > 1$, it degenerates into $\mathbb{R}_+ \times \mathbb{R}$. The bound H_ℓ is effective when $a \in]0, 1[$. When $a \geq 2$, although H_ℓ is not necessarily degenerate, it is less tight than the bound V_ℓ studied in the next section. It is only when $a \in]1, 2[$ that the inclusion of H_ℓ in Z_k provides some original information if $\ell \in]ok[$.

The bounds H_ℓ and H_∞ are of special interest if the transportation cost function $h(\delta)$ is a positive-coefficient linear combination of δ with expressions like δ^a (with $a \in]0, 1[$), or $\ln(\delta)$, or $-\delta^{-a}$ (with $a > 0$). The transportation cost function is then concave and $H_\infty \subseteq Z_k \subseteq H_\ell$, and neither of the bounds is a priori degenerate. On the other hand, an illustration of Prop. 4.9b is given by the case of $h(\delta) = \delta - \sqrt{(\delta + K)}$ with $K \geq 1/4$: then $H_\ell \subseteq Z_k \subseteq H_\infty$, with no a priori degeneracy of the bounds.

A question which we have not raised so far is whether the boundary of H_∞ is a limiting curve of $Z_j \cap Z_k$. As that boundary of H_∞ has itself two straight asymptotes, the problem boils down to the determination of whether those

straight asymptotes are limiting curves of $Z_j \cap Z_k$; we shall come back to this issue in Sections 4.8 and 4.10.

4.6. Asymptotic direction

Suppose that Z_k is not bounded and let a point i move towards infinity on $Z_j \cap Z_k$ above the x -axis. If the angle φ_i , under which i is seen from some fixed point i' of the plane tends towards some limit that we denote by φ_∞ , the angle $\varphi_{i''}$ must tend towards the same limit for any point i'' of the plane. We call that angle φ_∞ the *asymptotic direction* of $Z_j \cap Z_k$. When φ_∞ exists, it is easily seen (Fig. 4.4) that

$$\lim \{\delta_j - \delta_k : \delta_k \rightarrow \infty, i \in Z_j \cap Z_k\} = \delta_{jk} \cos (\varphi_\infty) .$$

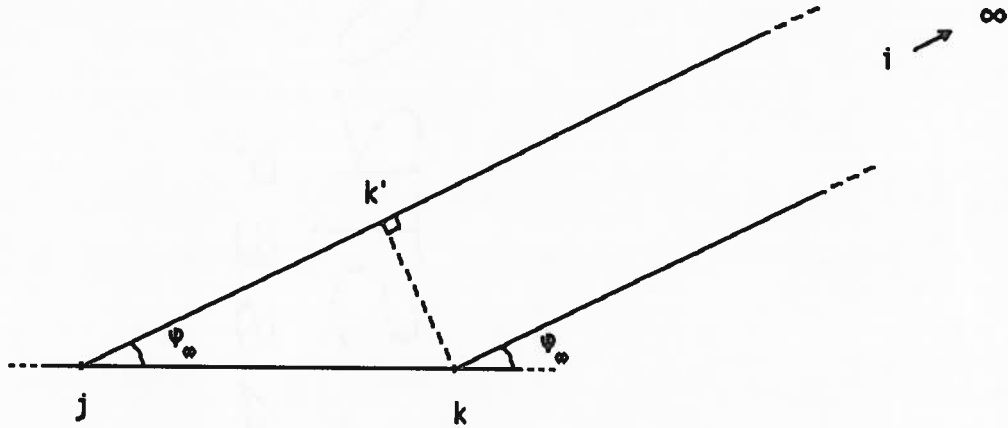


Fig. 4.4. — When i moves toward infinity, $\delta_j - \delta_k \rightarrow \delta_{jk} \cos (\varphi_\infty)$.

As that limit is also equal to $Q/h'(\infty)$ (see the proof of Prop. 4.2), we finally get

$$\cos (\varphi_\infty) = Q / h'(\infty) \delta_{jk} . \quad (4.3)$$

Conversely, when Z_k is unbounded and $h'(\infty)$ exists, $Z_j \cap Z_k$ has the asymptotic direction given by (4.3).

It can be seen in particular that the boundary of H_∞ (Section 4.5) has the same asymptotic direction as $Z_j \cap Z_k$. Another interesting point is that the property ' $0 \leq \cos(\varphi_\infty) \leq 1$ ' (as $\varphi_\infty \in [0, \pi/2]$) is perfectly coherent with the conditions of unboundedness of Z_k ; see Prop. 4.2. Formula (4.3) is also in accordance with Prop. 2.5: as Z_k is a decreasing function of the attractivity constant Q , so must be the angle φ_∞ .

4.7. Vertical bounds and vertical limiting lines

Assuming that $h''(\infty)$ exists lato sensu, we define the following vertical half-planes:

$$V_\ell = \{i: x \geq x_\ell\}$$

$$V_\infty = \{i: x \geq Q / \delta_{jk} h''(\infty)\} .$$

Of those two sets, the only one that may degenerate is V_∞ . This happens when $h''(\infty) = 0$ and when $h''(\infty) = \infty$; in the first case, we have $V_\infty = \emptyset$; in the second, $V_\infty = \mathcal{R}_+ \times \mathcal{R}$. Notice that $h''(\infty) \geq 0$, as will appear from the proof of Prop. 4.10. Last, here again, in order to simplify the formulation, when $h''(\infty) \notin \mathcal{R} \cup \{-\infty, \infty\}$, we set $V_\infty = \emptyset$ if $h \circ \sqrt{\cdot}$ is concave, and $V_\infty = \mathcal{R}^2$ if $h \circ \sqrt{\cdot}$ is convex. Notice that Prop. 4.10 may appear as a consequence of Prop. 3.12, as far as the vertical bound V_ℓ is concerned.

(4.10) (a) If $h''(\infty) \in \mathcal{R}_+^*$, the boundary of V_∞ is a limiting line of $Z_j \cap Z_k$;

(b) If $h''(\infty) = 0$, $Z_j \cap Z_k$ has no vertical limiting line;

(c) If $h \circ \sqrt{\cdot}$ is concave, $V_\infty \subseteq Z_k \subseteq V_\ell$;

(d) If $h \circ \sqrt{\cdot}$ is convex, $V_\ell \subseteq Z_k \subseteq V_\infty$;

(e) If $h \circ \sqrt{\cdot}$ is concave or convex and $h''(\infty) \in \mathcal{R}_+^* \cup \{\infty\}$,

the boundary of V_∞ is asymptotic to $Z_j \cap Z_k$.

The proof relies again on the mean-value theorem, now applied to the function $h \circ \sqrt{\cdot}$ (which is derivable if h is). From this and from equality (2.4) we deduce that

$$\begin{aligned} \Delta h &= h(\sqrt{\delta_j^2}) - h(\sqrt{\delta_k^2}) \\ &= (h \circ \sqrt{\cdot})'(\delta^2) (\delta_j^2 - \delta_k^2) = [h'(\delta)/\delta] \cdot x \cdot \delta_{jk} \end{aligned} \quad (4.4)$$

for some $\delta \in]\delta_k, \delta_j[$. If $h''(\infty) \in [-\infty, \infty]$, we have according to de l'Hospital's rule, even if $h'(\infty) \neq \infty$,

$$\lim [h'(\delta)/\delta: \delta \rightarrow \infty] = h''(\infty) .$$

Consequently, $h''(\infty)$ is ≥ 0 as announced, and (4.4) entails that, if $h''(\infty) > 0$,

$$x_\infty = \lim [x: \delta_k \rightarrow \infty, i \in Z_j \cap Z_k] = Q / h''(\infty) \delta_{jk} . \quad (4.5)$$

As ' $h''(\infty) > 0$ ' implies that $h'(\infty) = \infty$ (see de l'Hospital's rule), Z_k is unbounded and the limit (4.5) makes sense. Hence the first item. Conversely, if there is a vertical limiting line, Z_k is unbounded and we may use (4.5), which shows that $h''(\infty) \neq 0$. This proves item (b) by contraposition. As to the third item: the inclusion of Z_k in V_ℓ is a mere application of Prop. 3.5c; and, as $h \circ \sqrt{\cdot}$ is concave, we have

$$(h \circ \sqrt{\cdot})'(\delta^2) = h'(\delta) / 2\delta \geq (h \circ \sqrt{\cdot})'(\infty) = h''(\infty) / 2 ,$$

so that if $i \in V_\infty$, then $\Delta h \geq h''(\infty) \cdot x \cdot \delta_{jk} \geq Q$ because of (4.4), and $i \in Z_k$. The argument is similar for item (d). And item (e) obviously results from items (c) and (d), from the proof of item (a), and from Prop. 2.3. **Q.E.D.**

In the case of the monomial transportation cost function δ^a , Prop. 4.10 shows that $Z_j \cap Z_k$ has no vertical limiting curve when $a \in [1, 2[$, and that the axis of y is the straight asymptote of $Z_j \cap Z_k$ when $a > 2$ (see Fig. 3.4). When the transportation cost is for instance a positive-coefficient linear combination of δ^2 with expressions like $\ln(\delta)$ or δ^a with $a \in]0, 2[$, we face a more complex situation where a vertical asymptote exists which is not the y -axis.

Notice also that the property 4.10c, concerning V_{ℓ} , has a consequence that we express in these familiar terms: when h is concave, $Z_j \cap Z_k$ turns its concavity in ℓ toward k .

4.8. Oblique limiting lines

Suppose now that $Q \in]\delta_{jk} h'(\infty), \infty[$, so that Z_k is unbounded and $Z_j \cap Z_k$ has a nonvertical and nonhorizontal asymptotic direction: $\varphi_{\infty} \notin \{0, \pi/2\}$. The question here is: is some straight line $y = \alpha x + \beta$ a limiting line of $Z_j \cap Z_k$? Of course, as far as the upper part of the plane is concerned, the coefficient α must be equal to $\operatorname{tg}(\varphi_{\infty})$; and $\beta = \lim [y - \alpha x: i \in Z_j \cap Z_k, \delta_k \rightarrow \infty]$. We have

$$\begin{aligned} y - x \operatorname{tg}(\varphi_{\infty}) &= \frac{y^2 - x^2 \operatorname{tg}^2(\varphi_{\infty})}{y + x \operatorname{tg}(\varphi_{\infty})} \\ &= \frac{\delta_o^2 - x^2 [1 + \operatorname{tg}^2(\varphi_{\infty})]}{y + x \operatorname{tg}(\varphi_{\infty})} \\ &= \frac{\delta_o^2 \cos^2(\varphi_{\infty}) - x^2}{x \cos^2(\varphi_{\infty}) [\operatorname{tg}(\varphi_o) + \operatorname{tg}(\varphi_{\infty})]} . \end{aligned}$$

We now express x and δ_o^2 as functions of δ_k and the difference $\delta_j - \delta_k$, in short $\Delta\delta$. As $x = (\delta_j^2 - \delta_k^2) / 2\delta_{jk}$ and $\delta_o^2 = (\delta_j^2 + \delta_k^2)/2 - \delta_{ok}^2$, this yields

$$x = \Delta\delta(2\delta_k + \Delta\delta) / 2\delta_{jk}$$

and

$$\delta_o^2 = \delta_k^2 + \delta_k \Delta\delta + \frac{(\Delta\delta)^2}{2} - \delta_{ok}^2 ,$$

so that

$$\delta_o^2 \cos^2(\varphi_{\infty}) - x^2 = \delta_k (\delta_k + \Delta\delta) \left[\cos^2(\varphi_{\infty}) - \frac{(\Delta\delta)^2}{\delta_{jk}^2} \right] + C ,$$

where

$$C = \left[\frac{(\Delta\delta)^2}{2} - \delta_{ok}^2 \right] \cos^2(\varphi_{\infty}) - \frac{(\Delta\delta)^4}{4\delta_{jk}^2} .$$

As $\Delta\delta \in]0, \delta_{jk}[$, C is bounded above and C/δ_o tends towards 0 when $\delta_k \rightarrow \infty$. Let us now consider the difference $\Delta\delta/\delta_{jk} - \cos(\varphi_{\infty})$. According to the mean-

value theorem, if h is derivable on $]\delta_k, \delta_j[$, $\Delta h = h'(\delta^*) \Delta \delta = Q$ for some $\delta^* \in]\delta_k, \delta_j[$; taking also (4.3) into account, we have:

$$\cos(\varphi_\infty) - \frac{\Delta \delta}{\delta_{jk}} = \frac{Q}{\delta_{jk}} \left[\frac{1}{h'(\infty)} - \frac{1}{h'(\delta^*)} \right].$$

Let us now define $f(1/\delta) = 1/h'(\delta)$, so that, if that function f is derivable on $]0, 1/\delta^*[$ (i.e., if, on $]\delta^*, \infty[$, h is twice derivable and h' does not take 0 as value), and using again the mean-value theorem, we find for some $\delta^- > \delta^*$:

$$\begin{aligned} \cos(\varphi_\infty) - \Delta \delta / \delta_{jk} &= Q \frac{f(0) - f(1/\delta^*)}{\delta_{jk}} \\ &= - \frac{Q f'(1/\delta^-)}{\delta^* \delta_{jk}} \\ &= - \frac{Q \delta^{-2} h''(\delta^-)}{\delta^* \delta_{jk} h'^2(\delta^-)}. \end{aligned} \quad (4.6)$$

Hence,

$$\delta_0^2 \cos^2(\varphi_\infty) - x^2 = - \frac{\delta_k (\delta_k + \Delta \delta) \left[\frac{\Delta \delta}{\delta_{jk}} + \cos(\varphi_\infty) \right] Q \delta^{-2} h''(\delta^-)}{\delta^* \delta_{jk} h'^2(\delta^-)} + C,$$

so that we get, noticing that $x = \delta_0 \cos(\varphi_0)$ and that $\Delta \delta / \delta_{jk} \rightarrow \cos(\varphi_\infty)$ when $\delta_k \rightarrow \infty$:

$$\begin{aligned} \beta &= \lim [y - \alpha x: \delta_k \rightarrow \infty, i \in Z_j \cap Z_k] \\ &= - \frac{Q \lim \left[\frac{\delta_k^2 \delta^{-2} h''(\delta^-)}{\delta^* \delta_0}: \delta_k \rightarrow \infty, i \in Z_j \cap Z_k \right]}{\delta_{jk} h'^2(\infty) \sin(\varphi_\infty) \cos(\varphi_\infty)}. \end{aligned}$$

As δ_k/δ^* and δ_k/δ_0 tend towards 1, we finally have, given (4.3):

$$\beta = - \frac{\lim [\delta^2 h''(\delta): \delta \rightarrow \infty]}{h'(\infty) \sin(\varphi_\infty)}. \quad (4.7)$$

When β exists, it is clear that the conditions allowing to derive that formula are satisfied: the existence of $h'(\infty)$ and the fact that $h'(\infty) \neq 0$ imply that, when δ is above some $\underline{\delta}$, h is derivable and h' does not take 0 as value; the fact that $\lim [\delta^2 h''(\delta): \delta \rightarrow \infty] \in \mathfrak{R}$ implies that $\underline{\delta}$ may be chosen large enough for h to be twice derivable above $\underline{\delta}$.

Formula (4.7) can be used, for instance, when the transportation cost $h(\delta)$ is a positive-coefficient linear combination of δ with $\ln(\delta)$, or with expressions of the type $-\delta^{-a}$ with $a > 0$. In such a case $Z_j \cap Z_k$ has two oblique limiting lines, which pass through point o iff. $\ln(\delta)$ is none of the functions composing the linear combination; see Fig. 4.2 and 4.5.

We may also combine what we know about hyperbolic bounds with (4.7); this yields

(4.11) When $h'(\infty) \in]Q/\delta_{jk}, \infty[$, the half-hyperbola $\equiv \delta_j - \delta_k = Q/h'(\infty)$ is a limiting curve of $Z_j \cap Z_k$ iff. $\lim [\delta^2 h''(\delta) : \delta \rightarrow \infty] = 0$.

If, moreover,

(a) h is concave or convex, that half-hyperbola is asymptotic to $Z_j \cap Z_k$;

(b) h is convex, the oblique lines $\equiv |y| = x \operatorname{tg}(\varphi_\infty)$ are asymptotes of $Z_j \cap Z_k$.

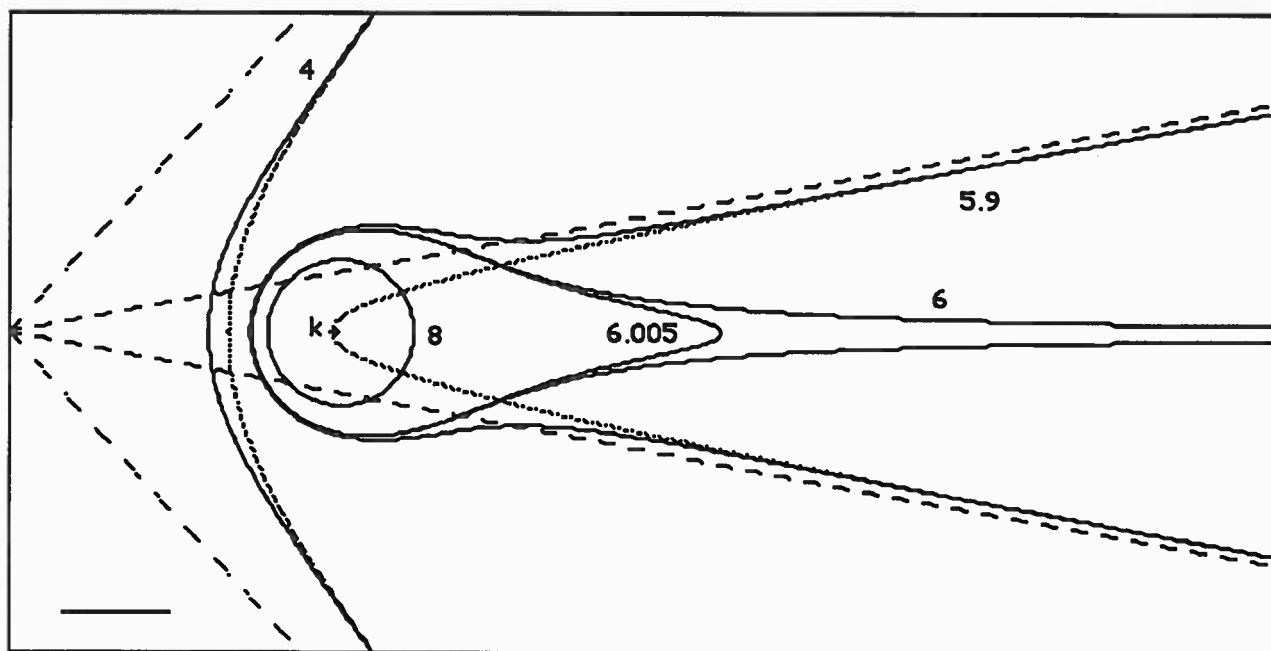


Fig. 4.5. — Market areas generated by the t.c.f. $\delta - (\delta + 0.1)^{-4}$ when $\delta_{jk} = 6$, for indicated values of Q . The x -axis is the horizontal asymptote of the indifference line when $Q = 6$; the superficies of Z_k is then finite, however. When $Q < 6$, the half-hyperbola $\equiv \delta_j - \delta_k = Q$ is asymptotic to the indifference line and contained in Z_k .

When $h(\delta) = \delta - \delta^{-4}$, for instance, $Z_j \cap Z_k$ has two oblique asymptotes $y = x \operatorname{tg}(\varphi_\infty)$ and $y = -x \operatorname{tg}(\varphi_\infty)$; the half-hyperbola $\equiv \delta_j - \delta_k = Q$ is asymptotic to the indifference line and contained in Z_k (Fig. 4.5; for the finiteness of the measure of Z_k when $Q = 6$, see Hanjoul *et al.*). Prop. 4.11 also trivially applies, of course, when $h(\delta) = \delta$.

4.9. Horizontal limiting lines

From (2.3) and (2.4), we compute that

$$y^2 = \frac{\delta_j^2 + \delta_k^2}{2} - \frac{(\delta_j^2 - \delta_k^2)^2}{4\delta_{jk}^2} - \delta_{ok}^2. \quad (4.8)$$

Replacing δ_j by $\delta_k + \Delta\delta$ in that expression, we get

$$y^2 = \delta_k^2 \left(1 + \frac{\Delta\delta}{\delta_k}\right) \left[1 - \frac{(\Delta\delta)^2}{\delta_{jk}^2}\right] - \delta_{ok}^2 \left[1 - \frac{(\Delta\delta)^2}{\delta_{jk}^2}\right]^2. \quad (4.9)$$

If $Z_j \cap Z_k$ has a horizontal asymptotic direction (i.e., if $\varphi_\infty = 0$), (4.3) implies that $Q = \delta_{jk} h'(\infty)$; setting $f(1/\delta^2) = 1/h'(\delta)$, and assuming that function f to be derivable on $]0, 1/\delta^{*2}[$, we then derive similarly to (4.6) that, for some $\delta^* \in]\delta_k, \delta_j[$ and some $\delta^- > \delta^*$, we have when point $i \in Z_j \cap Z_k$:

$$\begin{aligned} 1 - \frac{\Delta\delta}{\delta_{jk}} &= \frac{Q}{\delta_{jk}} \left[\frac{1}{h'(\infty)} - \frac{1}{h'(\delta^*)} \right] \\ &= - \frac{Q \delta^{-3} h''(\delta^-)}{2 \delta_{jk} \delta^{*2} h'^2(\delta^-)} \\ &= - \frac{\delta^{-3} h''(\delta^-) h'(\infty)}{2 \delta^{*2} h'^2(\delta^-)}. \end{aligned} \quad (4.10)$$

If we let δ_k tend toward infinity, $h'(\delta^-)$ tends toward $h'(\infty)$, $\Delta\delta \rightarrow \delta_{jk}$ (as $\varphi_\infty = 0$), and $\delta_k/\delta^* \rightarrow 1$. The limiting value of y^2 is, consequently,

$$\begin{aligned} y_\infty^2 &= \lim [y^2: \delta_k \rightarrow \infty, i \in Z_j \cap Z_k] \\ &= - \lim [\delta^3 h''(\delta): \delta \rightarrow \infty] / h'(\infty). \end{aligned} \quad (4.11)$$

Here again, it is easily seen that the existence of the limit implies that the hypotheses leading to that formula are satisfied.

According to (4.11), when the limit of $\delta^3 h''(\delta)$ is > 0 , y_∞^2 should be strictly negative: the conclusion is that Z_k must be bounded in such a case. This is again coherent with Prop. 4.2d, as the strict positivity of that limit entails the strict convexity of function h beyond some value $\underline{\delta}$ of its argument.

When $y_\infty = 0$, the horizontal limiting line, here reduced to the x -axis, is of course an asymptote of $Z_j \cap Z_k$:

(4.12) *When $Q = h'(\infty) \delta_{jk}$ and h is concave, the x -axis is asymptotic to $Z_j \cap Z_k$ iff. $\lim [\delta^3 h''(\delta) : \delta \rightarrow \infty] = 0$.*

Applying those results, for instance to transportation costs $h(\delta)$ that are positive-coefficient linear combinations of δ with expressions of the type $-\delta^{-a}$ with $a \geq 1$ shows that the indifference line $Z_j \cap Z_k$ is then endowed, at the threshold value $\delta_{jk} h'(\infty)$ of the attractivity constant Q , with a horizontal limiting line which is the x -axis unless one of those expressions is $-1/\delta$; see the examples of Fig. 4.5 and 4.6. When $Q < \delta_{jk} h'(\infty)$, we have seen that there are two oblique limiting lines with such transportation cost functions. This is also an illustration of the following property, which is obvious if one compares (4.7) and (4.11):

(4.13) *If $Z_j \cap Z_k$ has one horizontal limiting line or two when $Q = \delta_{jk} h'(\infty)$, then it has, when $Q < \delta_{jk} h'(\infty)$, two oblique limiting lines passing through point o .*

On the contrary, when $h(\delta) = \delta + \ln(\delta)$ for instance, there is no horizontal limiting line if $Q = \delta_{jk} h'(\infty)$, but there are two oblique ones if $Q < \delta_{jk} h'(\infty)$. Those oblique lines do not cross at o but at some point u of the x -axis on the left of o ; h is indeed concave, so that $x_u < 0$: see formula (4.7) and Fig. 4.2.

If we now include in the linear combinations considered above expressions of the type δ^a with $0 < a < 1$, there is neither oblique nor horizontal limiting line any more. Notice also that formulæ (4.7) and (4.11) and Prop. 4.11, 4.12, and 4.13 trivially apply when $h(\delta) = \delta$, as Z_k then shrinks to the straight half-line $[\delta_{ok}, \infty[\times \{0\}$ when $Q = \delta_{jk}$, and has two oblique asymptotes passing through o when $Q < \delta_{jk}$.

The knowledge thus acquired of the shape of market areas at the threshold value $\delta_{jk} h'(\infty)$ of Q , which marks the limit under which Z_k is unbounded and above which it is bounded, allows because of Prop. 2.5 to predict some features of the market areas when Q is close to $\delta_{jk} h'(\infty)$. This is particularly perceptible in the case of Fig. 4.5.

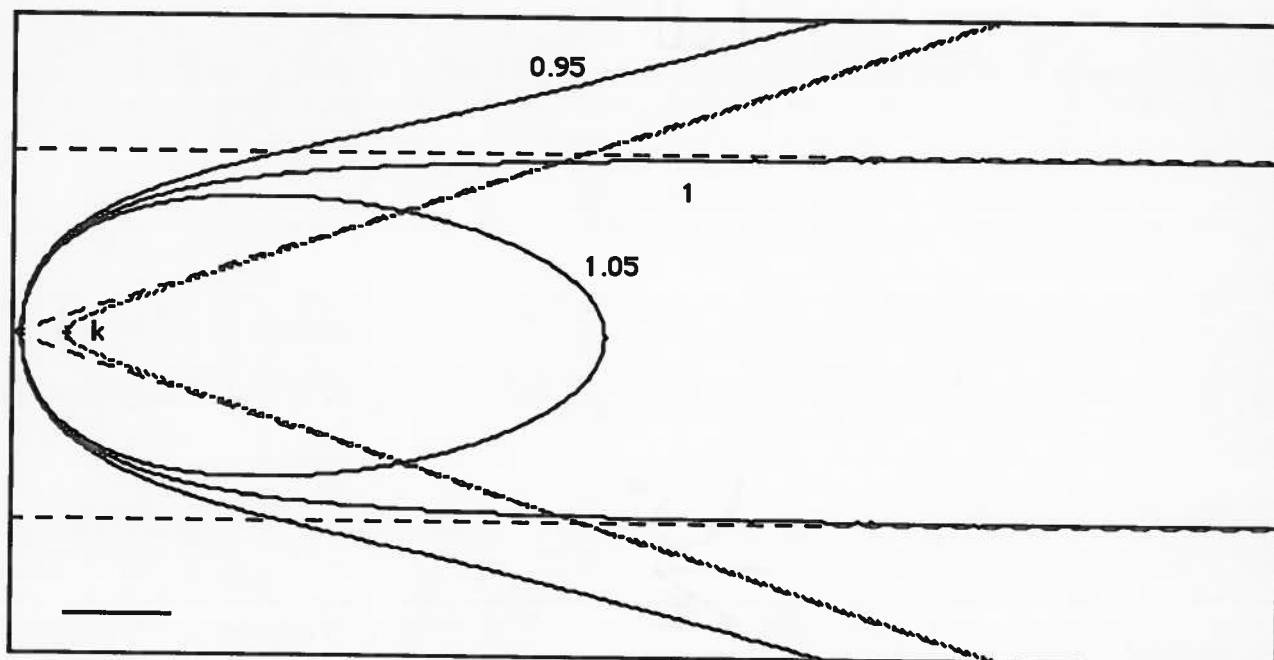


Fig. 4.6. — Market areas when the t.c.f. is $\delta - 1.5/\delta$ and $\delta_{jk} = 1$, for indicated values of Q . When $Q = 1$, the indifference line has two horizontal limiting lines $\equiv y = \pm \sqrt{3}$; when $Q < 1$, it is asymptotic to the half-hyperbola $\equiv \delta_j - \delta_k = Q$, which is contained in Z_k .

5. The standpoint of mathematical topology: spatial structure

The properties listed in this section are usually introduced in the mathematical field of general topology. Spinedness, however, seems a concept of our own, but we include it here because of its resemblance to starshapedness. Also, we have thought it interesting to mention visibility as another point of view on starshapedness. Mark that only two of the properties deserve the name of topological: these are arcwise-connectedness and simple connectedness. A property is topological iff. it is maintained through any homeomorphism; see e.g. Lipschutz (1965).

5.1. Arcwise-connectedness

A set E is *arcwise connected* (i.e., unipartite) iff. any two points of E can be joined by a *path* (i.e., the image set of some continuous application of a real interval onto the plane) contained in E . Otherwise said, E is unipartite (but may have holes: see Section 5.2). A maximal arcwise connected subset of a set E (be E unipartite or not) is an arcwise connected component of E . The arcwise connected components of E constitute a partition of E .

(5.1) (a) Z_j is arcwise connected;

(b) The unipartite components of $Z_k \cap [\delta_{0k}, \infty[\times \{0\}$ are the intersections of the unipartite components of Z_k with $[\delta_{0k}, \infty[\times \{0\}$, or equivalently the circular projections of the unipartite components of Z_k around k on $[\delta_{0k}, \infty[\times \{0\}$;

(c) Z_k has no more than one unbounded arcwise connected component;

(d) If h is concave or convex, Z_k is arcwise connected.

The items (b), (c), and (d) remain true if ' $Z_k - Z_j$ ' replaces ' Z_k '.

Item (d) applies for instance to the monomial transportation cost function (.)^a: Z_k is then arcwise connected anyway. Conversely, Fig. 5.1a displays a case where h is first concave, then convex. It can be shown, using Prop. 5.1b, that Z_k may then be bipartite as occurs here.

The proofs of Prop. 5.1a relies upon Prop. 3.1. This one entails that any point of Z_j either is in the half-plane $\mathcal{H}_x$ or can be joined to it by an arc Γ of a circle centred on any point $p \in [j_0]$, with $\Gamma \subseteq Z_j$. Any two points of Z_j may consequently be joined together by a path contained in Z_j , and item (a) is proved; see Fig. 5.2a.

Let us now come to the other items, for which we give the proof concerning Z_k only. As to item (b), also based on Prop. 3.1: of course, we mean by circular projection of a point i around k the intersection point of the circle centred on k and passing through i with the half-axis $[\delta_{0k}, \infty[\times \{0\}$. Let us denote that point by $\pi(i)$. The first step is to prove that, C being a unipartite component of Z_k , we have $\pi\langle C \rangle = C \cap [\delta_{0k}, \infty[\times \{0\}$. The argument is straight as $\pi\langle C \rangle \subseteq C \cap [\delta_{0k}, \infty[\times \{0\}$ (from Prop. 3.1) and, obviously,

$$C \cap [\delta_{0k}, \infty[\times \{0\} = \pi \langle C \cap [\delta_{0k}, \infty[\times \{0\} \rangle \subseteq \pi \langle C \rangle .$$

(Of course, we have also $Z_k \cap [\delta_{0k}, \infty[\times \{0\} = \pi \langle Z_k \rangle$, and this will ease our notation).

The second step is to show that when a set C is a unipartite component of Z_k , then $\pi\langle C \rangle$ is a unipartite component of $\pi\langle Z_k \rangle$; i.e., a maximal unipartite subset of $\pi\langle Z_k \rangle$. First, let us examine why $\pi\langle C \rangle$ is unipartite. Consider indeed two points i_1 and i_2 of $\pi\langle C \rangle$. According to Prop. 3.1, $\pi\langle C \rangle \subseteq C$; so i_1 and i_2 may be linked by a path $P \subseteq C$. But Prop. 3.1 again shows that $\pi\langle P \rangle$, which is obviously a path between i_1 and i_2 , is contained in $\pi\langle C \rangle$. So $\pi\langle C \rangle$ is unipartite. Were it not maximal, $\pi\langle C \rangle$ would be strictly included in a broader unipartite subset of $[\delta_{0k}, \infty[\times \{0\}$. We could thus find a point $i' \in \pi\langle Z_k \rangle$, not contained in $\pi\langle C \rangle$ but linkable to it by some path P' included in $\pi\langle Z_k \rangle$. Being outside $\pi\langle C \rangle$, i' would also be outside C , as $\pi\langle C \rangle = C \cap [\delta_{0k}, \infty[\times \{0\}$; P' being included in Z_k , C would not be a maximal unipartite component of Z_k : a contradiction. (To be exhaustive we ought also to prove that if C' is a unipartite component of $\pi\langle Z_k \rangle$, then $C' = \pi\langle C \rangle$ for some unipartite component C of Z_k ; but this is rather obvious).

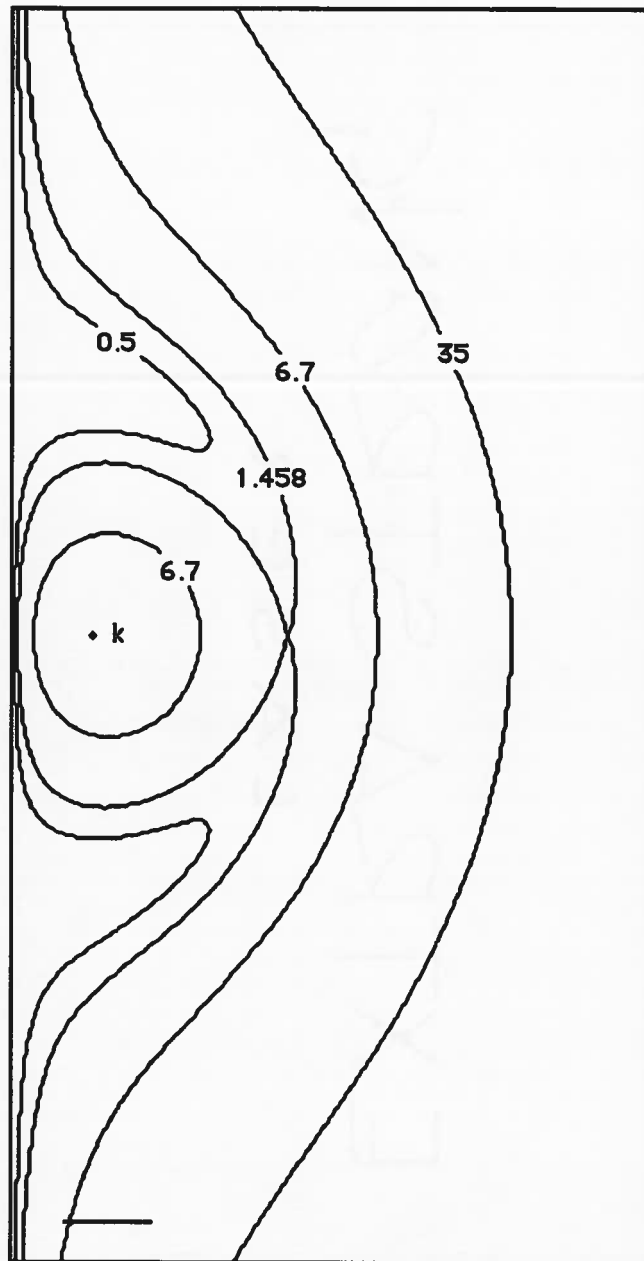


Fig. 5.1 a. — Market areas produced by a t.c.f. first concave, then convex:

$$h(\delta) = (\delta - 3.1)^3, \delta_{jk} = 1.8, \text{ indicated values of } Q.$$

Item (c) directly derives from item (b): as a subset of a straight half-line, $\pi\langle Z_k \rangle$ must have no more than one unipartite component. As we know from Prop. 4.2a the equivalence ' C bounded $\Leftrightarrow \pi\langle C \rangle$ bounded', the proof is complete.

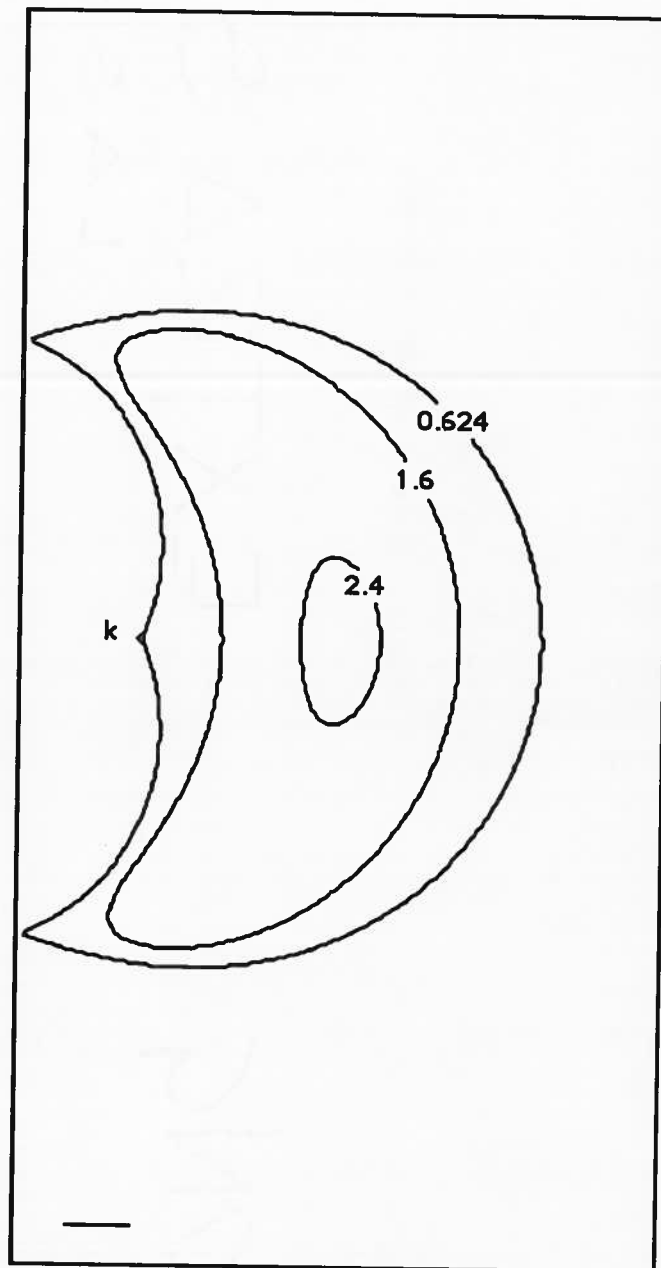


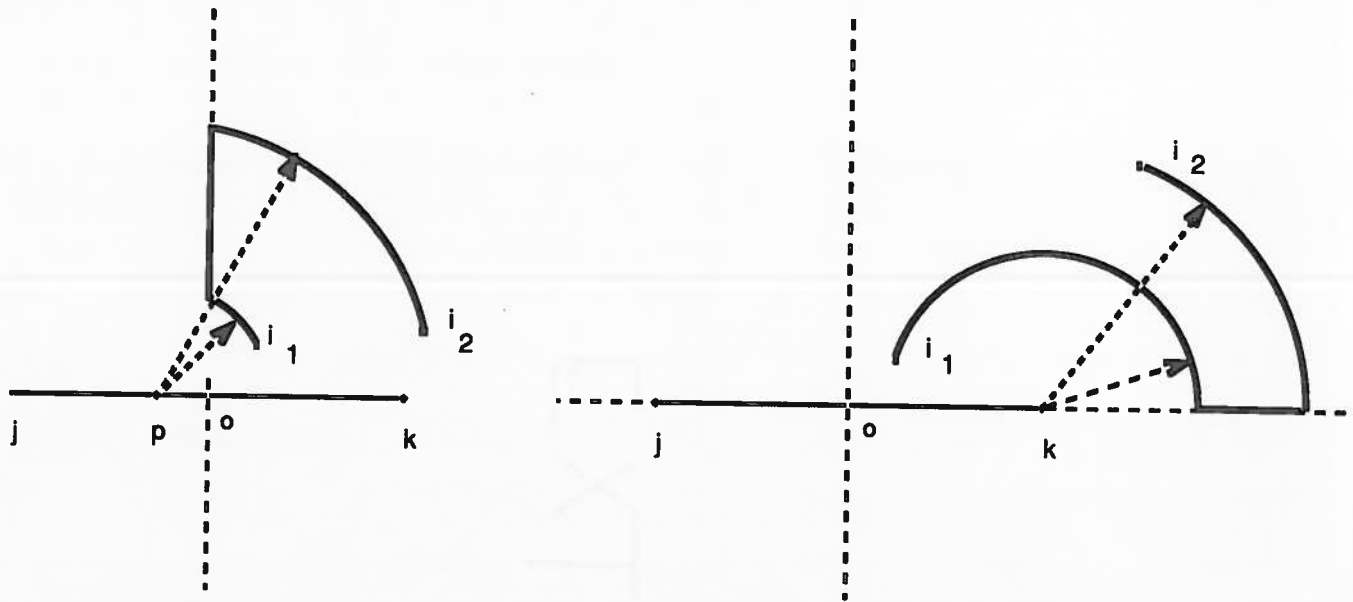
Fig. 5.1 b. — Market areas produced by a t.c.f. first convex, then concave:

$$h(\delta) = (\delta - 4.8)^{1/3}, \delta_{jk} = 3.6, \text{ indicated values of } Q.$$

Item (d) also is a mere application of item (b). We have seen (cf. the proof of Prop. 4.1b and 4.1c) that Δh is a monotonic function of x on $[\delta_{ok}, \infty[\times \{0\}$ when h is concave or convex. Consequently, if i_1 and i_2 belong to $\pi\langle Z_k \rangle$, we have for any $i \in [i_1 i_2]$:

$$\Delta h \geq \min \{ \Delta h_{i_1}, \Delta h_{i_2} \} \geq Q$$

so that $[i_1 i_2] \subseteq \pi\langle Z_k \rangle$. So $\pi\langle Z_k \rangle$ and Z_k are unipartite; see Fig. 5.2b. Q.E.D.



a) Z_j is arcwise connected.

b) Z_k is arcwise connected when h is concave or convex.

Fig. 5.2. — Arcwise-connectedness.

5.2. Simple connectedness

A set E is *simply connected* iff. every loop (or closed path; i.e., a path in which, starting from a point i , we come back to it at the end of the path) of E is homotopic through E to a point of E . In $\mathbb{R}^2$, that means the impossibility of drawing in E a loop that would surround a point of $\mathbb{R}^2 - E$. Otherwise said, E has no hole, it has no Emmenthal-like structure. (Some authors additionally require that E should be arcwise connected, but when the above definition is used E may independently be arcwise or simply connected). The two properties are related, however:

LEMMA 5.1. If $\mathbb{R}^2 - E$ is arcwise connected and unbounded, E is simply connected.

Indeed, were E not simply connected, it would be possible to draw in E a path around some $i \in \mathbb{R}^2 - E$. If $\mathbb{R}^2 - E$ is arcwise connected, it should consequently be bounded. Q.E.D.

Concerning market areas, we get the following properties:

- (5.2) (a) Z_k is simply connected;
- (b) Z_j [resp. $Z_j - Z_k$] is simply connected iff. $Z_k - Z_j$ [resp. Z_k] is unbounded and arcwise connected, or empty;
- (c) If h is convex, or if $Q \in [\Delta h_k, h'(\infty) \delta_{jk}]$ and h is concave, Z_j is simply connected.

We give the proofs about the simple connectedness of Z_k and Z_j only; that concerning $Z_j - Z_k$ is perfectly similar. Item (a) is a consequence Lemma 5.1, as $Z_j - Z_k$ is arcwise connected (same argument as with Prop. 5.1a) and of course unbounded. Concerning (b) and (c), notice first that Z_j is simply connected when $Z_k - Z_j$ is empty; and that the condition for this, when h is concave, is that $\Delta h_k \leq Q$. Now assume that $Z_k - Z_j \neq \emptyset$. The implication ($\Leftarrow$) of item (b) is identical to Lemma 5.1. As to ($\Rightarrow$): suppose that $Z_k - Z_j$ is bounded or not unipartite. In the first case, it is possible to surround $Z_k - Z_j$ with a path included in Z_j , and Z_j is not simply connected. In the second, $(Z_k - Z_j) \cap (\mathbb{R}_+ \times \{0\})$ is not unipartite (Prop. 5.1b). So we can find on $\mathbb{R}_+ \times \{0\}$ two points, i_1 and i_2 , of $Z_k - Z_j$ separated by a point $i' \in Z_j$, with $x_{i_1} < x_{i_2}$. According to Prop. 3.1, the circle passing through i' and centred on any point p of $[j_0]$ is then contained in Z_j ; as it surrounds $i_1 \in Z_k - Z_j$, market area Z_j is not simply connected. Finally, item (c) is derived from Prop. 5.2b, 5.1d, and 4.2b. **Q.E.D.**

Simple connectedness thus appears as another viewpoint on boundedness and arcwise-connectedness. It is easily seen, in particular, that Z_j is not simply connected when $h = (.)^a$ with $0 < a < 1$, provided that $Z_k - Z_j \neq \emptyset$, but that Z_j is simply connected when $a \geq 1$.

As to the distinction between Z_k and $Z_k - Z_j$ or between Z_j and $Z_j - Z_k$, notice that if a set E is simply connected, then the suppression of points of E cannot make possible to draw new paths inside E . That is why we did not mention in item (a) that $Z_k - Z_j$ is simply connected. But $Z_j - Z_k$ may be simply connected without Z_j possessing the same property. If we lower the attractivity constant Q in the example of Fig.5.1a, the two unipartite

components $Z_k^{(1)}$ and $Z_k^{(2)}$ will grow. When $Q = 1.458$, $Z_k^{(1)}$ and $Z_k^{(2)}$ will share a unique point lying on the x -axis; so that there will be in fact one unipartite component $Z_k^{(1)} \cup Z_k^{(2)}$ for Z_k , instead of two, $Z_k^{(1)} - Z_j$ and $Z_k^{(2)} - Z_j$, for $Z_k - Z_j$. In that case $Z_j - Z_k$ is simply connected while Z_j is not. But the distinction is of course more mathematical than geographical ...

5.3. Spinedness

Let us set this definition: a set E is *spined* by a straight line Δ iff. for every $i \in E$, the segment $[ii']$ is a subset of E , i' being the orthogonal projection of i on Δ .

It follows from Prop. 3.2 and 3.3 that:

- (5.3) (a) The part of Z_k lying at the left of k , i.e., $Z_k \cap [0, \delta_{ok}] \times \mathcal{R}$, is spined by the straight line perpendicular to $[jk]$ and passing through k , i.e., $\{\delta_{ok}\} \times \mathcal{R}$;
- (b) If h is convex, Z_j is spined by the y -axis;
- (c) If $h \circ \sqrt{\cdot}$ is concave, Z_k is spined by the x -axis;
- (d) If $h \circ \sqrt{\cdot}$ is convex, Z_j is spined by the x -axis.

The property of spinedness is not independent of the ones studied so far. In particular, it is easily seen that if a set $E \subseteq \mathcal{R}^2$ is spined by a straight line Δ , then E is simply connected and, if $\mathcal{R}^2 - E$ is moreover contained in one of the two half-planes delimited by Δ , $\mathcal{R}^2 - E$ is also simply connected.

According to that it is possible to give, for instance, a new proof of the simple connectedness of Z_j and Z_k when function h is convex. Noticing that Prop. 5.3 may be extended to the interior of the concerned areas, and using Prop. 5.3b, we can indeed build those two logical chains:

$$h \text{ convex} \Rightarrow Z_j \text{ spined} \Rightarrow Z_j \text{ simply connected};$$

$$\begin{array}{l|l} h \text{ convex} \Rightarrow Z_j - Z_k \text{ spined by the } y\text{-axis} & \\ Z_k \subseteq \mathbb{R}_+ \times \mathbb{R} & \Rightarrow Z_k \text{ simply connected.} \end{array}$$

The property is also related to the property of arcwise-connectedness: when h is convex and $h \circ \sqrt{\cdot}$ is convex or concave [as is the case if $h = (\cdot)^a$ with $a \geq 1$], Prop. 5.3b, c, and d entail that any two points i_1 and i_2 of Z_k (resp. Z_j) may be linked by a Manhattan-like path contained in Z_k (resp. Z_j). Last, Prop. 5.3b also gives a straightforward proof that when h is convex, $Z_k \neq \emptyset$ iff. Z_k is unbounded (Prop. 4.1c).

In the case of the monomial transportation cost function $(\cdot)^a$, Prop. 5.3 entails that Z_k is spined by the x -axis when $a \in]0, 2]$; and that Z_j is spined by the y -axis when $a \geq 1$, and by the x -axis when $a \geq 2$. On the contrary, in the two examples of Fig. 5.1, neither Z_k nor Z_j are spined by the axes of x or y ; (a) is the only item that can be used (if we do not consider separately the properties of the unipartite components of Z_k , as far as Fig. 5.1a is concerned).

Fig.5.3 shows possible shapes of market areas when function h is made of linear pieces. In Fig. 5.3a, h is concave; in Fig. 5.3b, it is convex. Though surprising, those shapes are coherent with Prop. 5.3. In the first case h is concave but not convex: no wonder if Z_k is not spined by the y -axis. In the second case, neither Z_j nor Z_k are spined by the x -axis; and indeed $h \circ \sqrt{\cdot}$ is neither convex nor concave.

5.4. Starshapedness and visibility

A set E is *starshaped* w.r.t. a point i iff., $\forall p \in E$, the segment $[ip]$ is contained in E . In more common terms: E is *radial* w.r.t. i , or: E *radiates* from i . The set E is *starshaped* w.r.t. to a set S iff. it is starshaped w.r.t. $\forall i \in S$.

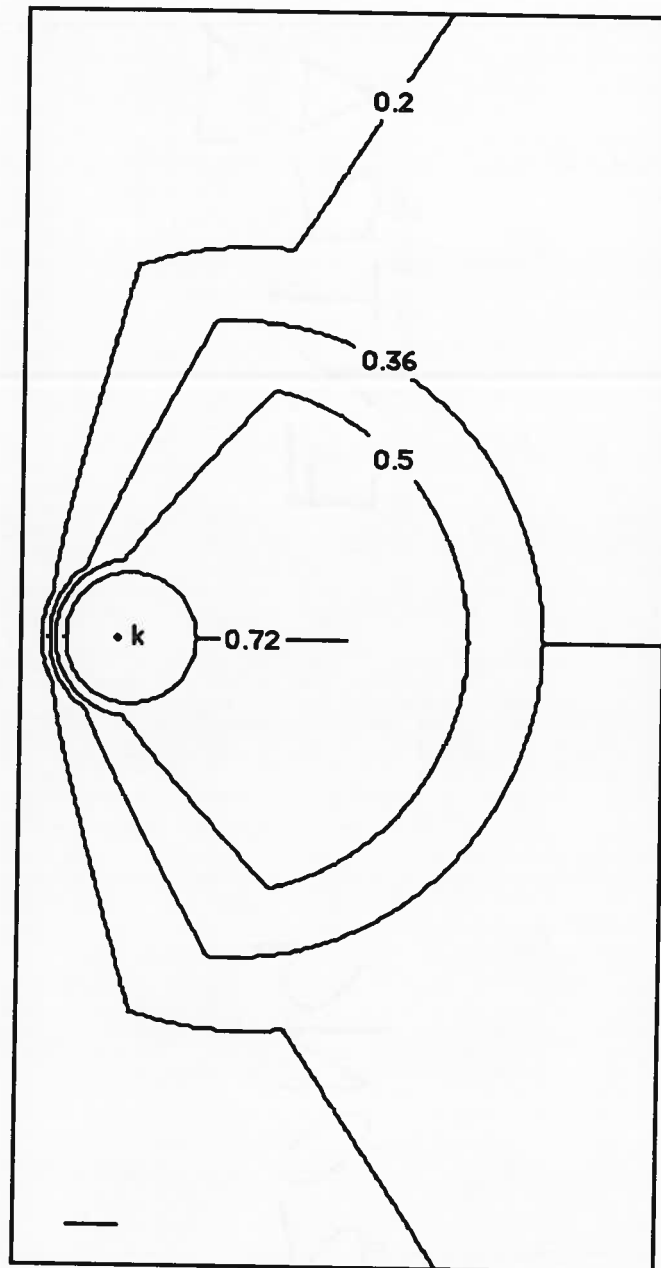


Fig. 5.3a. — $h(\delta) = \min \{\delta, 0.2\delta + 1.2, 0.1\delta + 2\}$, $\delta_{jk} = 3.6$, indicated values of Q .
Notice the infinite 1-dimensional tail of Z_k when $Q = 0.36$ and the finite one when $Q = 0.72$.

- (5.4) (a) If h is concave, Z_k is starshaped w.r.t. k ;
 (b) If h is convex, Z_j is starshaped w.r.t. $\mathcal{H}_- \times \{0\}$, the nonpositive part of the x -axis;
 (c) $Z_k \cap D(o, k)$ is starshaped w.r.t. k , and $Z_j \cap D(o, k)$ is starshaped w.r.t. j .

The proof is straightforward. Item (a) and (c) are just another formulation of Prop.3.2a and 3.2c. Similarly, Prop. 3.2b yields that Z_j is starshaped w.r.t. o (and even w.r.t. $[jo]$) when h is convex, and item b is obtained by combining this with Prop. 3.3b. **Q.E.D.**

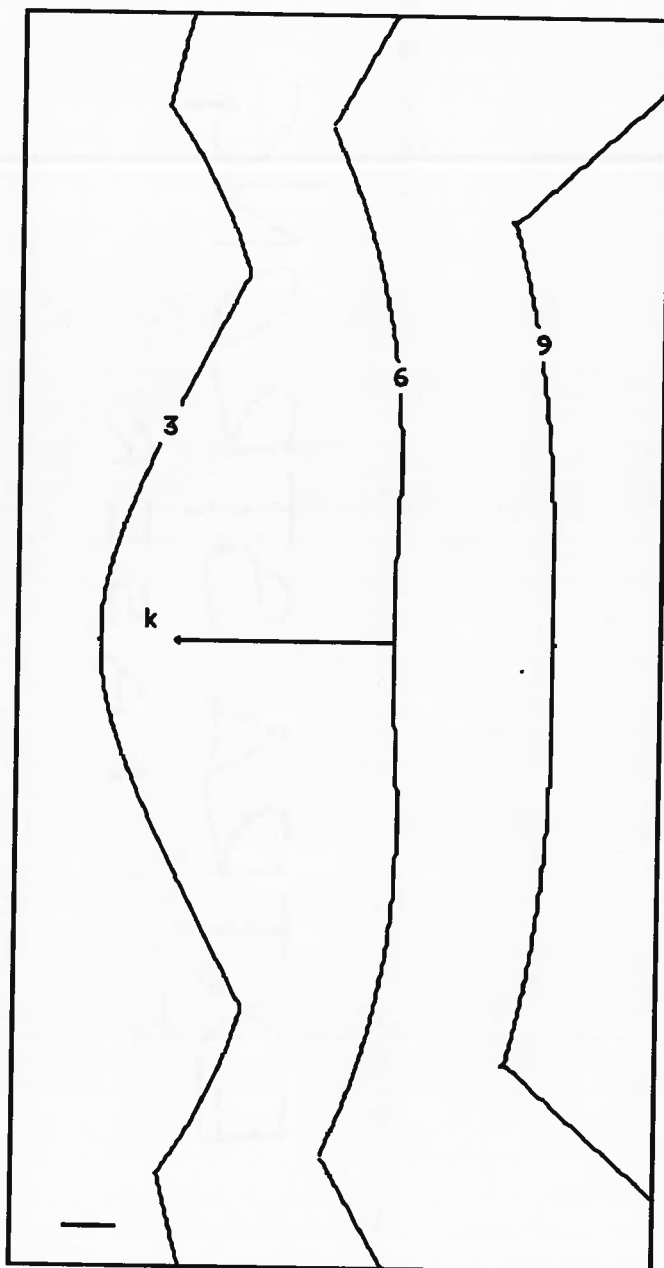


Fig 5.3b. $h(\delta) = \max \{2\delta, \delta + 10\}$, $\delta_{jk} = 6$, indicated values of Q .
Notice that, when $Q = 6$, Z_k has a one-dimensional tail ending at k .

As was the case for spinedness, starshapedness pours a new light on preceding properties. It is immediately seen that if a set E is starshaped w.r.t. a point p , E is arcwise connected, as $\{i_1, i_2\} \subset E \Rightarrow [i_1p] \cup [pi_2] \subset E$

E. The set E is also simply connected: otherwise some point $i \in \mathbb{R}^2 - E$ would be surrounded by a closed path $\Gamma \subseteq E$; calling i' an intersection point of Γ with the straight half-line originating in i and the prolongation of which passes through p , we should have $[pi'] \subseteq E$ and hence $i \in E$, a contradiction. If moreover E is unbounded, $\mathbb{R}^2 - E$ is simply connected because of Lemma 5.1. The propositions 5.1d (when function h is concave) and 5.2c may consequently be considered as deriving from Prop. 5.4. The Propositions 4.1b and 4.1c about the possible emptiness of market area Z_k are also corollaries of Prop. 5.4.

Prop. 5.4 can be illustrated by most of the examples given so far. See in particular Fig. 3.3, 5.1 and 5.3.

A concept very close to starshapedness is that of visibility, introduced by Goldman (1963) in the field of convex programming and by Hurter et al. (1975) in location theory. Adapting slightly their definitions, we say that a set E is *entirely visible* from a point p iff., for all $i \in E$, no other point of E is on the segment $]ip]$; i.e., $]ip] \cap E = \emptyset$. It is *entirely visible* from a set S iff. it is entirely visible from every $p \in S$. So we have concerning the border $Z_j \cap Z_k$ a proposition parallel to Prop. 5.4:

(5.5) (a) *If h is strictly concave [resp. strictly convex], $Z_j \cap Z_k$ is entirely visible from k [resp. $\mathbb{R}_- \times \{0\}];$*

(b) *If h is concave [resp. convex], $(Z_j \cap Z_k) -]\delta_{ok}, \infty[\times \{0\}$ is entirely visible from k [resp. $\mathbb{R}_- \times \{0\}].$*

The restriction expressed in item (b) corresponds to the possibility for $Z_j \cap Z_k$ to have a nondenumerable intersection with the straight half-line $] \delta_{ok}, \infty[\times \{0\}$ (see e.g. Prop. 4.2d).

A remarkable case occurs when the t.c.f. is linear: $h(\delta) = \delta$. That function being concave and convex, both items of Prop. 5.4 (or 5.5) can be applied. As a consequence, each market area is starshaped w.r.t. its centre. That property obtains whatever the number of centres; see Section 7.1.

5.5. Convexity

This mathematical concept corresponds to the intuitive idea of a round shape. A set E is *convex* iff., for any two points i_1 and i_2 of E , the segment $[i_1 i_2]$ is in E . The set E is *strictly convex* iff. $\forall i_1, i_2 \in E$, the open segment $]i_1 i_2[$ is in E° , the interior of E ; i.e., E minus its boundary. In other words, E is convex iff. it is starshaped w.r.t. itself. A convex set E is thus arcwise and simply connected; it is also spined by any straight line containing a diameter of E , a diameter being understood here as a segment $[i' i'']$ for which i' and i'' belong to E and maximize the distance $\delta_{i', i''}$ on E .

Convexity is thus a strong property. It is then perhaps not surprising that our results in this matter lack the polarization observed in the preceding statements between the concavity and convexity of h or of some derived function; the items (c) and (d) of Prop. 5.6 even concern only particular classes of transportation cost functions. To ease subsequent developments, we define from now on that a t.c.f. h is *convexifying* iff., when two centres j and k are concerned, the market area Z_k is always convex provided that $Q \geq 0$ ($p_k \geq p_j$); h is *strictly convexifying* iff., in the same circumstances, Z_k is strictly convex unless Z_k is contained in the x -axis or $Q = 0$.

- (5.6) (a) Except when $Z_j = \mathbb{R}^2$, Z_j is convex iff. $Z_j \cap Z_k$ is a straight line parallel to the y -axis; Z_j is strictly convex iff. $Z_j = \mathbb{R}^2$;
- (b) If the function $\sqrt{\cdot}/(h' \circ \sqrt{\cdot})$ is derivable, convex and strictly increasing, or if it is derivable, concave and strictly decreasing on the range of squared distances corresponding to $Z_j \cap Z_k$, Z_k is strictly convex;
- (c) When $h = (\cdot)^a$,
- (i) if $0 < a < 1$, Z_k is strictly convex;
 - (ii) if $a = 1$, Z_k is convex; if moreover $Q \neq \delta_{jk}$, Z_k is strictly convex;
 - (iii) if $1 < a < 2$ or $a > 2$, Z_k is not convex;
 - (iv) if $a = 2$, Z_k is convex but not strictly;

(d) When $h = \ln [K + (.)^a]$ with $K > 0$ and $a = 1$ or 2 , Z_k is strictly convex.

Item (a) is obvious as $\mathfrak{R}_- \times \mathfrak{R} \subset Z_j$. Item (b) is a direct consequence of Prop. 4.4; or of formula (3.15) as $\Delta g'$ and $\Delta h'$ have opposite signs. Item (d) derives from item (b) when $a = 2$, and from Hanjoul et al., Prop. 13.7a.

As to item (c): when $a = 1$, $Z_j \cap Z_k$ is a branch of a hyperbola, and Z_k is consequently strictly convex, unless it is a straight half-line ($Z_k^0 = \emptyset$). When $a > 1$, we know from (4.3) that $\phi_\infty = \pi/2$: exactly as in item (a) regarding Z_j , this implies, if Z_k is convex and not empty, that $Z_j \cap Z_k$ is vertical. As this is true only if $a = 2$, and as $Z_k \neq \emptyset$ if $a > 1$ [because $h'(\infty) = \infty$: see Prop. 4.1c and 4.2b], the proof of item (c) is complete, except for (c.i). That last point is proved by using (3.12) and (3.13) to show that, on $\mathfrak{R}_+^{*2}$, when $h = (.)^a$ with $0 < a < 1$, $(\partial^2 y / \partial x^2)_{\Delta h}$ is < 0 and finite. Those heavy computations form the remainder of the proof.

Let $\xi = -\Sigma h' / \Delta h'$, $\rho = \delta_j / \delta_k$, and $\lambda = \delta_{jk}^2 / \delta_k^2$. Then (3.12) rewrites:

$$\left(\frac{\partial y}{\partial x}\right)_{\Delta h} = \frac{\frac{\xi \delta_{jk}}{2} - \frac{\delta_j^2 - \delta_k^2}{2\delta_{jk}}}{y} = \delta_{jk} \frac{\xi \lambda - (\rho^2 - 1)}{2y\lambda}. \quad (5.1)$$

On the other hand, from (2.3), we have:

$$\begin{aligned} y^2 &= \frac{2(\delta_{jk}^2 \delta_j^2 + \delta_{jk}^2 \delta_k^2 + \delta_j^2 \delta_k^2) - \delta_j^4 - \delta_k^4 - \delta_{jk}^4}{4\delta_{jk}^2} \\ &= \delta_k^4 \frac{2(\lambda \rho^2 + \lambda + \rho^2) - \rho^4 - 1 - \lambda^2}{4\delta_{jk}^2} \\ &= \delta_{jk}^2 \frac{-\lambda^2 + 2\lambda(\rho^2 + 1) - (\rho^2 - 1)^2}{4\lambda^2}. \end{aligned} \quad (5.2)$$

Combining both results, we obtain:

$$\begin{aligned} 1 + \left(\frac{\partial y}{\partial x}\right)_{\Delta h}^2 &= 1 + \frac{[\xi \lambda - (\rho^2 - 1)]^2}{-\lambda^2 + 2\lambda(\rho^2 + 1) - (\rho^2 - 1)^2} \\ &= \lambda \frac{(\xi^2 - 1)\lambda + 2[\rho^2(1 - \xi) + 1 + \xi]}{-\lambda^2 + 2\lambda(\rho^2 + 1) - (\rho^2 - 1)^2}. \end{aligned}$$

Formula (3.13) has shown that $y(\partial^2 y / \partial x^2)_{\Delta h}$ is the difference between $(\partial^2 y^2 / \partial x^2)_{\Delta h} / 2 + 1$ and that quantity. According to (3.14),

$$\frac{1}{2} \left(\frac{\partial^2 y^2}{\partial x^2} \right)_{\Delta h} + 1 = \lambda \zeta \quad \text{where} \quad \zeta = \frac{4\delta_{jk}^2 \bar{h}'_k{}^2 \bar{h}'_j{}^2 \Delta g'}{\Delta \bar{h}'^3} . \quad (5.3)$$

Hence,

$$y \left(\frac{\partial^2 y}{\partial x^2} \right)_{\Delta h} = -\lambda \frac{\lambda^2 \zeta - \lambda [2\zeta(\rho^2+1) - \xi^2 + 1] + [\zeta(\rho^2-1)^2 + 2\rho^2(1-\xi) + 2 + 2\xi]}{-\lambda^2 + 2\lambda(\rho^2 + 1) - (\rho^2 - 1)^2} .$$

When $h = (\cdot)^a$, we have

$$h'(\delta^2) = \frac{a}{2} (\delta^2)^{(a/2)-1} = \frac{a}{2} \delta^{a-2}$$

$$h''(\delta^2) = \frac{a}{2} \left(\frac{a}{2} - 1 \right) (\delta^2)^{(a/2)-2} = \frac{a}{4} (a - 2) \delta^{a-4}$$

$$g'(\delta^2) = (-h''/2h')(\delta^2) = \frac{2-a}{2a} \delta^{-a} .$$

We so find that

$$\xi = \frac{\rho^{2-a} + 1}{\rho^{2-a} - 1}$$

$$\zeta = (2 - a) \rho^{2-a} \frac{\rho^a - 1}{(\rho^{2-a} - 1)^3} .$$

Introducing this in the expression of $y(\partial^2 y / \partial x^2)_{\Delta h}$ here above, we find that

$$y \left(\frac{\partial^2 y}{\partial x^2} \right)_{\Delta h} = \frac{\lambda \rho^{2-2a} T(\lambda, \rho)}{(\rho^{2-a} - 1)^3 [-\lambda^2 + 2\lambda(\rho^2 + 1) - (\rho^2 - 1)^2]} , \quad (5.4)$$

where

$$T(\lambda, \rho) = -\lambda^2(2-a)(\rho^a-1) + 2\lambda[(4-a)\rho^a + (2-a)\rho^{2+a} - (4-a)\rho^2 - (2-a)] + C$$

and where C is an expression of ρ (and of a , but a is here a constant number), independent of λ : $(\partial C / \partial \lambda)_\rho = 0$. From the fact that $\delta_j > \delta_k$ on $\mathfrak{X}_+^* \times \mathfrak{X}$ and from (5.2), it is clear that $(\partial^2 y / \partial x^2)_{\Delta h}$ and $T(\lambda, \rho)$ have the same sign at every point i of $\mathfrak{X}_+^{*2}$. When ρ is fixed, point i is on an Apollonian circle. The minimum of λ on that curve occurs when δ_k is maximum; i.e., when $\lambda = (\delta_j - \delta_k)^2 / \delta_k^2 = (\rho - 1)^2$. Conversely, λ is maximum when $\lambda = (\delta_j + \delta_k)^2 / \delta_k^2 =$

$(\rho + 1)^2$. We have thus to study T on the range $[(\rho - 1)^2, (\rho + 1)^2]$ of λ . Now, at the beginning of that interval, we compute that:

$$\left(\frac{\partial T(\lambda, \rho)}{\partial \lambda}\right)_\rho [(\rho - 1)^2, \rho] = 4\rho^a E(\rho),$$

where

$$E(\rho) = -\rho^{2-a} - (2-a)\rho^{1-a} + (2-a)\rho + 1.$$

Let us compute the derivative E' :

$$E'(\rho) = (2-a)[1 - \rho^{1-a} - (1-a)\rho^{-a}],$$

which is $< 0 \forall \rho > 1$, as $a \in]0, 1[$. As $E(1) = 0$, we thus have $E(\rho) < 0$, $\forall \rho > 1$: which means that $(\partial T / \partial \lambda)_\rho$ is < 0 when $\lambda = (\rho - 1)^2$. But $T(\lambda, \rho)$ itself is concave w.r.t. λ ; $(\partial T / \partial \lambda)_\rho$ is thus < 0 for any $\lambda \geq (\rho - 1)^2$. Now, when $\lambda \rightarrow (\rho - 1)^2$, the point i on the Apollonian circle determined by the still fixed value of ρ becomes close to the x -axis; (5.1) then shows that $|(\partial y / \partial x)_{\Delta h}| \rightarrow \infty$ [it is clear that $\xi > 1$, for h is here strictly concave; so that $\xi\lambda - (\rho^2 - 1) \neq 0$ in (5.1) when $\lambda = \rho^2 - 1$], whereas we see from (5.3) that $1 + (\partial^2 y^2 / \partial x^2)_{\Delta h} / 2$ remains finite. Consequently, $y(\partial^2 y / \partial x^2)_{\Delta h}$ and *a fortiori* $(\partial^2 y / \partial x^2)_{\Delta h}$ tend towards $-\infty$: (5.2) and (5.4) then entail that $T[(\rho - 1)^2, \rho] < 0$. As we have proved above that $(\partial T / \partial \lambda)_\rho < 0$ when $\lambda \geq (\rho - 1)^2$, we thus conclude that $T(\lambda, \rho) < 0$, and that $(\partial^2 y / \partial x^2)_{\Delta h} < 0$. Which in turn implies the convexity of Z_k . **Q.E.D.**

Still about items (c.i) and (d), two facts are worth noticing. The first is that the function $\sqrt{.} / (h' \circ \sqrt{.})$ is here concave and strictly increasing: item (b) cannot be used to prove item (c); and indeed, as shown in Fig. 3.3 (and 13.3b of Hanjoul et al.), Z_k is characterized by $L_y / L_x \leq 1$, in accordance with Corollary 4.1a. The second fact is that the convexity of Z_k does not here derive from the possible concavity of Δh seen as a function of (x, y) : Prop. 5.6c entails that Δh is quasi-concave w.r.t. (x, y) ; it is not concave, however. To show this we just have to consider the evolution of Δh along the positive part of the x -axis, for instance: Δh is convex and increasing w.r.t. x on $[ok]$, then convex and decreasing on $[ok) - [ok]$.

We already know that item (a) is illustrated by $h = (.)^2$; and we have seen how Prop. 2.1 can then be used to find all the transportation cost functions for which $Z_j \cap Z_k$ is vertical. Item (b) applies for instance to the

unbounded unipartite component $Z_k^{(1)}$ of market area Z_k in Fig. 5.1a, as it can be computed that the required condition is verified inside the range of distances to j and k that concerns Z_k ; we have thus also here $L_y/L_x > 1$. Item (b) also applies when $h = -(\cdot)^{-a}$ with $a > 0$, and when $h = \ln [K + (\cdot)^2]$ with $K > 0$.

Notice, conversely, that we have met situations where function h is concave but where Z_k is obviously not convex: see Fig. 5.3a and 4.5. In the first case, the nonconvexity of Z_k is clearly related with the nonderivability of h at some given distance. In the second, however, function h is perfectly smooth. On the other hand, when h is partially convex and concave, Z_k may be convex, as indicated by Prop. 2.7e (Fig. 2.5), or not; see Fig. 2.2 and 5.1. All this shows that the convexity of Z_k is absolutely independent of the concavity of function h .

5.6. About interiors, boundaries, and closures

Let us first recall some definitions. The interior of a set E , E° , is the largest open set contained in E . A point p adheres to E iff., $\forall \varepsilon > 0$, $D_p(\varepsilon) \cap E \neq \emptyset$. The closure of E , $cl(E)$, may be defined as the set of all points adhering to E . The boundary of E , $b(E)$, may be defined as the set of all points adhering to E and to $\mathcal{R}^2 - E$ at the same time; hence, $b(E) = b(\mathcal{R}^2 - E)$. It is well-known that

$$\begin{aligned} b(E) \cap E^\circ &= \emptyset \\ cl(E) &= b(E) \cup E^\circ. \end{aligned} \tag{5.5}$$

We have already seen in Fig. 5.3 that some seemingly obvious properties concerning those sets are not necessarily verified in the case of Z_k and Z_j . It is clearly not always true, for instance, that $\mathcal{R}^2 - Z_k = Z_j^\circ$. We only know, *a priori*, that $\mathcal{R}^2 - Z_k \subseteq Z_j^\circ$, because $\mathcal{R}^2 - Z_k$ is an open set. Here are our main results; some of them will be useful in Section 8.2. From now on, to simplify our writings, we set that, for any set E , $E^- = \mathcal{R}^2 - E$.

$$(5.7) \quad (a) \quad Z_j^\circ \cap Z_k^\circ = \emptyset;$$

$$(b) \quad \text{cl}(Z_j^\circ) = Z_j;$$

$$(c) \quad (Z_j \cap Z_k) - ([\delta_{ok}, \infty[\times \{0\}) = b(Z_j) - ([\delta_{ok}, \infty[\times \{0\});$$

$$(d) \quad Z_j \cap Z_k = b(Z_k);$$

$$(e) \quad Z_j^- = Z_k^\circ;$$

$$(f) \quad \text{If } Z_k \text{ is convex and } Z_k^\circ \neq \emptyset, \quad Z_k^- = Z_j^\circ.$$

Those properties also hold when $Q = 0$.

Suppose that (a) is not true. Then there exists some $i \in Z_j^\circ \cap Z_k^\circ$. This has two consequences: $\Delta h_i = Q$, and $D_i(\varepsilon) \subseteq Z_j^\circ \cap Z_k^\circ$ for some $\varepsilon > 0$. In the upper intersection point i' of the boundary of that disk with $D(j, i)$, we have $\delta_{i',j} = \delta_{ij}$ and $\delta_{i',k} > \delta_{ik}$; hence, $\Delta h_{i'} < Q$ and $i' \notin Z_k$: a contradiction with the fact that $i' \in Z_j^\circ \cap Z_k^\circ$. Item (a) thus holds true.

Concerning item (b): as $Z_j \neq \emptyset$, take any point $i \in Z_j$. If $x_i \leq 0$, then $i \in \mathcal{R}_- \times \mathcal{R} \subseteq Z_j$ and $i \in \text{cl}(Z_j^\circ)$. If $x_i > 0$, the area $D(j, i) - D^\circ(k, i)$ is $\subseteq Z_j$. Obviously, $i \in \text{cl}([D(k, i) - D^\circ(j, i)]^\circ)$, which is contained in $\text{cl}(Z_j^\circ)$; we thus also have $i \in \text{cl}(Z_j^\circ)$. Hence, $Z_j \subseteq \text{cl}(Z_j^\circ)$; as $\text{cl}(Z_j^\circ) \subseteq \text{cl}(Z_j) = Z_j$ (since Z_j is a closed set), we finally have $Z_j = \text{cl}(Z_j^\circ)$.

For item (c), consider a point $i \in Z_j \cap Z_k$, not on $[\delta_{ok}, \infty[\times \{0\}$, the part of the x -axis lying on the right of k . That point thus adheres to Z_j ; and we have $\Delta h_i = Q$. As i is not on $[\delta_{ok}, \infty[\times \{0\}$, it is possible to move from i to the right on the circle $C(k, i)$, thus strictly increasing Δh . We so see that point i adheres to Z_j^- . As it also adheres to Z_j , it belongs to $b(Z_j)$. As a consequence,

$$(Z_j \cap Z_k) - ([\delta_{ok}, \infty[\times \{0\}) \subseteq b(Z_j) - ([\delta_{ok}, \infty[\times \{0\}) .$$

On the other hand, it is clear, because of the definition of boundaries, that $\Delta h = Q$ on $b(Z_j)$, because those points adhere to Z_j , where $\Delta h \leq Q$, and to Z_j^- , where $\Delta h > Q$. Hence, $b(Z_j) \subseteq Z_j \cap Z_k$. The rest of the proof of this item is obvious.

The proof of item (d) is very similar; but the point i is now simply taken out of $Z_j \cap Z_k$. It is also necessarily on the right of the y -axis, in $]0, \infty[\times \mathbb{R}$ (Prop. 2.3). It is thus possible to move from i to the left on the circle $C(j, i)$, thus strictly decreasing Δh . Hence, point i , which of course adheres to Z_k , also adheres to Z_k^- and belongs to $b(Z_k)$. We have so shown that $Z_j \cap Z_k \subseteq b(Z_k)$; as $b(Z_k) \subseteq Z_j \cap Z_k$ [see the proof of item (c)], we have the announced equality.

For item (e), take a point i in Z_k° . There is some $\varepsilon > 0$ for which $D_i(\varepsilon) \subseteq Z_k^\circ$. Hence, we may move to the left on $C(j, i)$ inside $D_i(\varepsilon)$. The points i' we meet are still in Z_k° , so that $\Delta h_{i'} \leq Q$. However, $\Delta h_{i'} < \Delta h_i$; thus, $\Delta h_i < Q$ and $i \in Z_j^-$. We so see that $Z_k^\circ \subseteq Z_j^-$. As the reverse inclusion is obvious (because Z_j^- is an open set contained in Z_k), we get the equality.

In item (f), we start similarly from a point $i \in Z_j^\circ$. Here again, for some $\varepsilon > 0$, we have $D_i(\varepsilon) \subseteq Z_j^\circ$. Suppose now that $i \in Z_k$ at the same time. As $Z_k^\circ \neq \emptyset$, there is some $i' \in Z_k^\circ$ and some ε' for which $D_{i'}(\varepsilon') \subseteq Z_k^\circ$. As Z_k is convex, the area $\cup\{[i''i'] : i'' \in D_{i'}(\varepsilon')\} \subseteq Z_k$. But the interior of that area is contained in Z_k° and obviously intersects $D_i(\varepsilon)$: hence, $Z_k^\circ \cap Z_j^\circ \neq \emptyset$, in opposition with item (a). Hence, $i \in Z_k^-$; here again, as it is clear that $Z_k^- \subseteq Z_j^\circ$, we have the equality $Z_k^- = Z_j^\circ$. **Q.E.D.**

An interesting immediate consequence of Prop. 5.7f is the following:

(5.8) *If h is convexifying and if $Z_k^\circ \neq \emptyset$, $Z_k^- = Z_j^\circ$ and $Z_j^- = Z_k^\circ$.*

5.7. Zero-dimensional market areas

The property of convexity has an interesting characteristic: it keeps an immediate meaning when the real set of customer locations, M , is finite. The convexity of a market area, Z_k for instance, implies indeed that the convex hull of the real market area of k , $Z_k \cap M$, is contained in Z_k and does not contain any point of the other real strict market area, $(Z_j - Z_k) \cap M$.

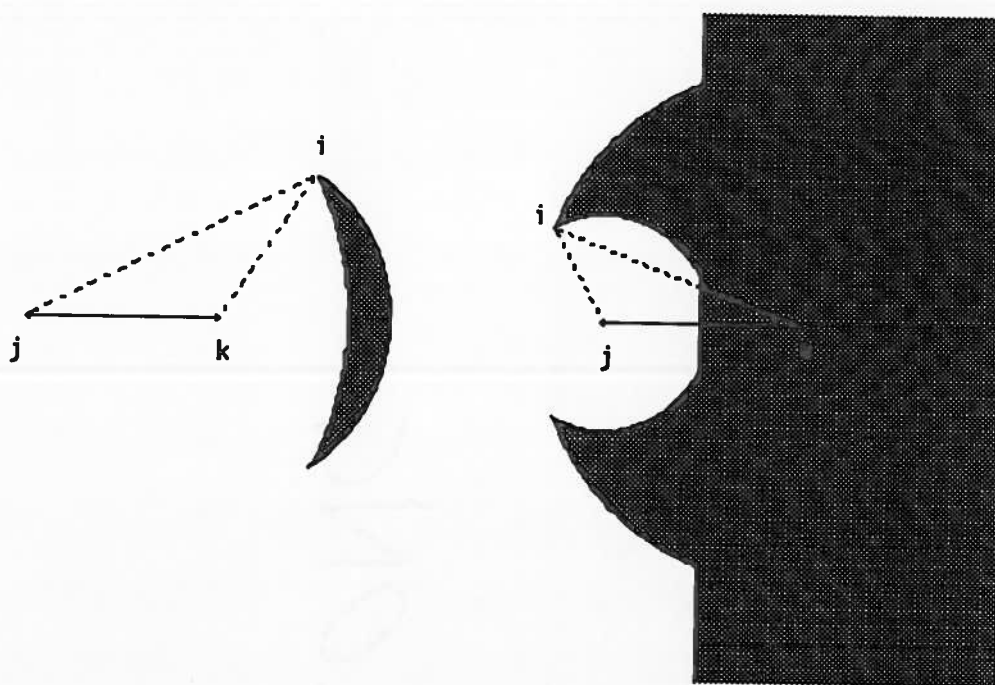


Fig. 5.4. — If a point i , not on the x -axis, belongs to a market area (e.g. Z_k), a corresponding two-dimensional area (B_{ik}) is contained in that market area.

The sign of Q is not supposed to be known *a priori*.

The other properties we have explored in this chapter all mean, however, that some curves (which are one-dimensional sets) are contained in market areas. This makes sense because those areas are two-dimensional. They are such because we consider the possibility of a nonzero mass m_i in every point i of the plane. We may be satisfied with such a description if we want to check if already established two-dimensional maps are in agreement with our theory. In the real world, however, the known customer locations are restricted to a finite set of relatively small areas distant from one another, cities for instance, or even to a finite set of points. It is thus difficult, if not impossible, to see if there is a contradiction between theory and reality: the probability that the above-mentioned curves meet the customer location set is small, or even zero when the set is finite, thus zero-dimensional. To increase that probability, we should replace those curves by two-dimensional sets which could intercept even zero-dimensional customer islets.

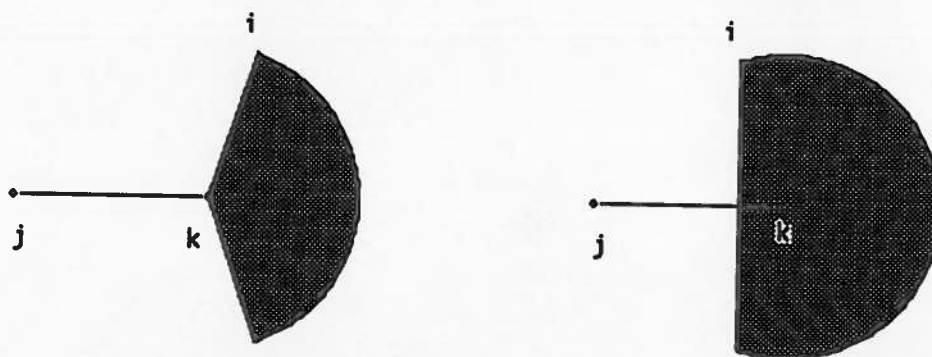


Fig. 5.5. — The area B_{ik} when h is concave, on basis of the present chapter ($Q > 0$, $x > 0$).

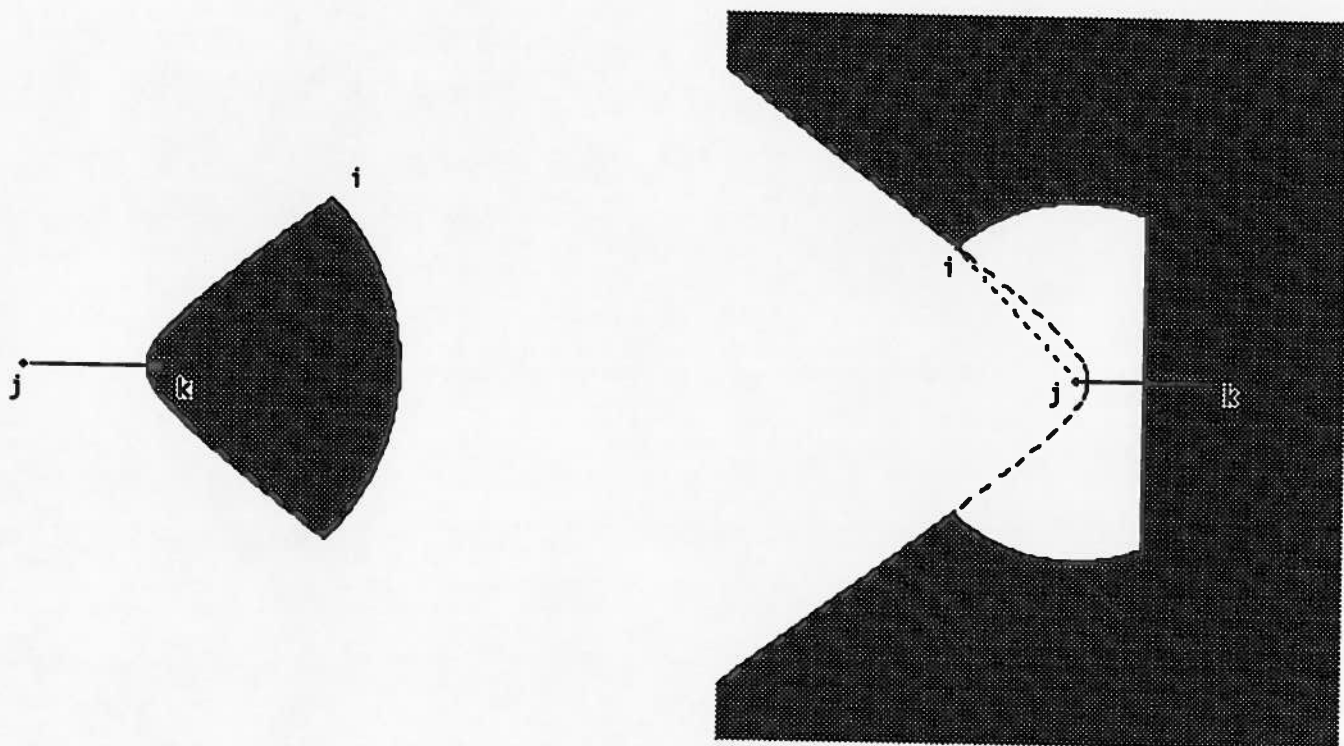


Fig. 5.6. — The best area B_{ik} when h is concave.

This is not a problem, hopefully. Suppose that we are initially at some location i (out of the x -axis) belonging for instance to Z_k ; it is clear we remain in Z_k if we transfer ourselves to another point i' in which $\delta_{i',j} \geq \delta_j$ and $\delta_{i',k} \leq \delta_k$. Otherwise said: $D(k, i) - D^\circ(j, i) \subseteq Z_k$ (see also Prop. 3.1). That area is two-dimensional and no point of $Z_j - Z_k$ should be found in it. It is depicted in Fig. 5.4, in which we also take into account the symmetry w.r.t. the x -axis and the fact that Q must be < 0 if some point $i \in Z_k \cap \mathcal{R}_-^* \times \mathcal{R}$. (All this remains true if we interchange j and k).

Spinedness and starshapedness thus keep a meaning: they may be combined with that property and so increase the two-dimensional areas that we know must be contained in Z_k or Z_j . Fig. 5.5 shows for instance those areas when h is concave (and $Q > 0$), on the basis of the propositions of the present chapter. However, in that particular case, the best result is obtained using (and developping) Prop. 3.4; the concerned curves are hyperbolæ having j and k as foci; see Fig. 5.6.

6. Generalizing the concept of market area

6.1. Individual comparison with other centres

We now assume that there are more than two centres. As announced in Chapter 1, the set of all centres is denoted by C . We adopt from now on a multiplicative formulation instead of the additive used in Part I. I.e., the comparison between centres is based on 'potential' [$V_c = a_c F(\delta_c)$], 'attractiveness', and 'distance deterrence function', rather than on 'delivered prices' [$p_c + h(\delta_c)$], 'mill prices', and 'transportation cost function'. If the initial model is an additive all-or-nothing one, we just have to set $a_c = \exp(-p_c)$ and $F(\delta) = \exp[-h(\delta)]$. The reason for that choice will appear from Section 6.2.

To simplify the writings, we introduce here some new symbols. The fact that j is more attractive than k (i.e., $a_j > a_k$, or $p_j < p_k$) will be denoted as $j \succ k$; $j \equiv k$ means that they are equally attractive, and $j \succeq k$ means that $j \succ k$ or $j \equiv k$. Similarly, $j \succ_i k$ means that j is preferred to k in location i (i.e., $V_j > V_k$, or $p_j + h_j < p_k + h_k$); $j \equiv_i k$ means that those quantities are equal, and $j \succeq_i k$ means that $j \succ_i k$ or $j \equiv_i k$.

Generally speaking, the area of a centre k may be expressed by the next formula:

$$V_k \geq U_k, \quad (6.1)$$

where U_k summarizes in some way the information about the other potentials, with the constraint that

$$U_{k\langle\{k, c\}\rangle} = V_c, \quad \forall c \in C - \{k\}.$$

In this formula and in others that will follow, we denote the value taken by some expression or set E when the set of centres, C , is replaced by some other, C' , as $E\langle C'\rangle$; here, $C' = \{k, c\}$.

A first possibility is to define U_k in this way:

$$U_k = \max \{V_c: c \in C - \{k\}\} \quad (6.2)$$

Equally valid from both the all-or-nothing and the distributive points of view, that formula defines what is usually considered a *market area* in the literature (possibly with other names, as we have said in Section 1.2). It is the area where $k \geq c$, $\forall c \in C$. We denote it as Z_k and maintain the name of market area. Its definition is thus:

$$Z_k = \{i: \forall c \in C: k \geq c\} = \{i: \forall c \in C: V_k \geq V_c\}. \quad (6.3)$$

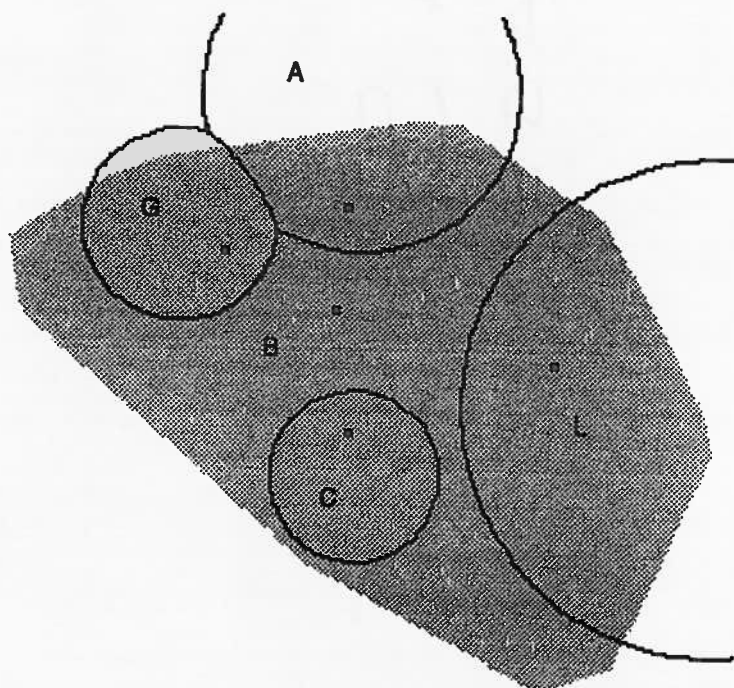


Fig. 6.1. — Market areas computed from Thill's model; see Fig. 10.4.

Fig. 6.1. illustrates this on the basis of a Pareto model calibrated by Thill [see Thill (1982, 1983) or Beguin and Thill (1985)] and based on empirical maps of the third level of urban hierarchy in Belgium; see Section 10.2. Attractivities are assumed to be proportional to populations, and the distance deterrence function is δ^{-a} . Thill found 3.2 as optimal value of the exponent a . We have represented Belgium in a simplified manner: the area in grey is a convex polygon very close to the convex hull of Belgium. Fig. 6.1 and similar ones have been drawn with the help of Mathematica 2.2.

.6.3.

Even at this early stage, obvious properties appears. First, the market areas are Voronoi cells w.r.t. $h(\delta)$ (see the definition of Voronoi cells w.r.t. δ in Section 7.1); every point of the plane thus belongs to one market area at least. Second, if we neglect the centres having an empty market area, we find of course the same pattern of market areas:

$$Z_k = Z_k \langle \{c: Z_c \neq \emptyset\} \rangle, \quad \forall k \in C.$$

In other words: when we read a map, *a posteriori*, we have not to take into account the centres having no market area to understand the shape of the market areas. Third, let us denote by G the following set:

$$G = \{c \in C: Z_c^\circ \neq \emptyset\}. \quad (6.4)$$

The other centres are of course not very interesting; and it is clear that every point of the plane belongs to some Z_c with $c \in G$. Moreover,

$$Z_k = Z_k \langle G \rangle, \quad \forall k \in G. \quad (6.5)$$

Even if we do not specify it, keep thus in mind that all the computations concerning a market area Z_k with nonempty interior may be made neglecting the centres belonging to $C-G$.

On the other hand, c and k being two centres, let us call Z_{kc} the market area that would be allocated to k if c were the only other centre:

$$Z_{kc} = Z_k \langle \{c, k\} \rangle = \{i: k \succeq_i c\}. \quad (6.6)$$

It is clear that we have here a third way to define the market area:

$$Z_k = \bigcap_{c \in C - \{k\}} Z_{kc}. \quad (6.7)$$

As a consequence, the boundary of Z_k is made of pieces of boundaries of the areas Z_{kc} and can be understood piecewise by means of the preceding sections.

In all-or-nothing models, Z_k is the area where the interaction t_{ik} is $\neq 0$ as soon as $d_i \neq 0$. It is thus usually the most suitable definition of the area

of k . Things are different, however, in distributive models (GG, MNL): the influence of k is generally not restricted to Z_k . In that case, particularly, we may think of other possible definitions of U_k . The following, for instance, is worth our attention:

$$U_k = \text{median } (V_c: j \in C - \{k\}) . \quad (6.8)$$

One may indeed choose as the indicator U_k the *median* of the other potentials (or the median of all potentials, but the two definitions are not equivalent). The median of a series of numbers $(V_1, V_2 \dots V_n)$, supposed to be ranked in decreasing order, is equal to $V_{(n+1)/2}$ when n is odd, $(V_{n/2} + V_{1+n/2})/2$ when n is even. We shall refer to those areas as *median areas*.

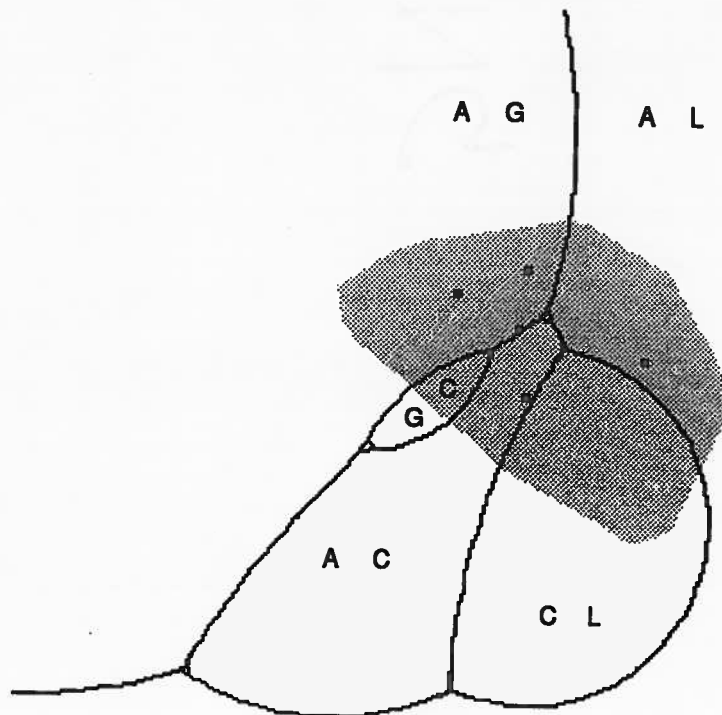


Fig. 6.2. — Median areas of the cities of the third level of the urban hierarchy in Belgium, based on Thill's model (comparison with all cities).

The median area of Brussels is the whole plane.

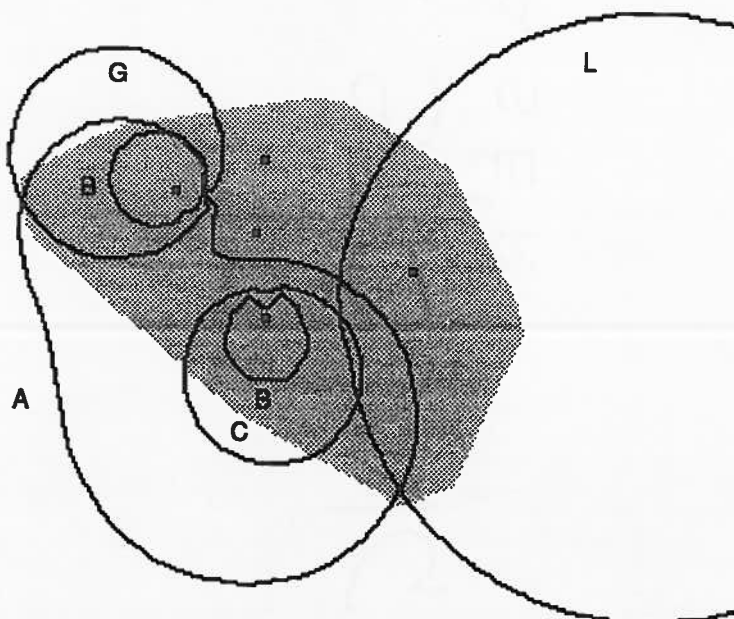


Fig. 6.3. — Median areas of the cities of the third level of the urban hierarchy in Belgium, based on Thill's model (comparison with the other centres).

In the same vein, we may also consider the area where k appears as the m^{th} choice among the centers, for any integer m between 2 and $\#C$ (when $m = 1$, the area is Z_k). Precisely, we say that

$$k \text{ is a } m^{\text{th}} \text{ choice for } i \Leftrightarrow \#\{c \in C: c \succ_i k\} = m - 1.$$

That kind of area is closely related to the concept of median: when n is odd, the area where V_k is $\geq$ to the median of all potentials is indeed also the area where k represents at most the $(n+1)/2^{\text{th}}$ choice. For instance, when there are five centres, if the potentials, ranked in decreasing order, appear as $V_{(1)}$, $V_{(2)}$, $V_{(3)}$, $V_{(4)}$, $V_{(5)}$, then V_k is $\geq$ to the median of all potentials when $V_k \in \{V_{(1)}, V_{(2)}, V_{(3)}\}$ and is thus the first, second, or third choice.

It so appears that one way to describe the influence of k is to consider the areas where k is the first or second choice; then the first, second, or third; etc. That kind of information may also be interesting in all-or-nothing models. We denote those areas by $Ch_k(1)$, $Ch_k(2)$, etc.:

.6.6.

$$\text{Ch}_k(m) = \{i: \#\{c: c \gg_i k\} < m\}.$$

Interestingly, this is equivalent to saying that

$$\text{Ch}_k(m) = \bigcup \{Z_k \langle C-P \rangle : P \subseteq C - \{k\} \text{ and } \#P = m-1\} . \quad (6.9)$$

For any integer $m > 1$, the m^{th} -choice area proper is of course equal to $\text{Ch}_k(m) - \text{Ch}_k(m-1)$. As is easily seen, $\text{Ch}_k(m)$ is a closed set, and a decreasing function of m . Of course, $\text{Ch}_k(1) = Z_k$ and $\text{Ch}_k(\#C) = \mathbb{R}^2$. Notice also that if U_k is chosen as

$$U_k = \min \{V_c : c \in C - \{k\}\} ,$$

the region allocated to k is $Ch_k(\#C-1)$.

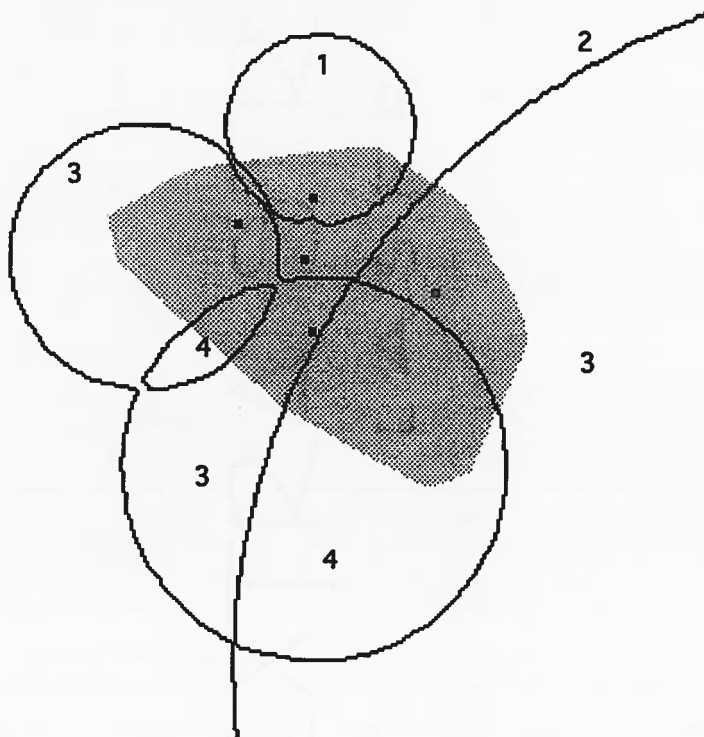


Fig. 6.4. — Choice areas of Antwerp, computed according to Thill's model.

The intersection of the disks respectively containing Gent and Antwerp is a part of the second-choice area.

We may also state an obvious proposition. We define that set $\mathcal{E}$ of areas is *chained* iff., for any A' and A'' of $\mathcal{E}$, there exists a series $(B_m: 1 \leq m \leq n, m \in \mathbb{N})$ of areas belonging to $\mathcal{E}$ for which $A' \cap B_1$, $A'' \cap B_n$, and $B_m \cap B_{m+1}$

.6.7.

($\forall m \in [1, m[\cap \mathbb{N})$ are $\neq \emptyset$. On the basis of (6.13), since the areas Z_C cover the whole plane, it is then clear that

(6.1) For any integer $m \geq 2$ and $\leq \#C$, the set of the areas $Ch_C(m)$ is chained.

Of course, that property also holds for median areas.

The notion of market area may still be extended in another way. Given some integer n between 2 and $\#C$, we define $Z_{c_1 c_2 \dots c_n}$ as the area where $c_1 \succeq_i c_2 \succeq_i \dots \succeq_i c_n$, for any point i of the area. We call such sets *preference areas*. The areas Z_{kc} appearing in (6.7) are an instance of this. Of course, we have

$$Z_{c_1 c_2 \dots c_n} = \bigcap (Z_{c_m c_{m+1}} : m \in \{1, 2, \dots, n-1\}) ; \quad (6.10)$$

and also, among other possible formulæ,

$$Z_{kj} = Z_{ckj} \cup Z_{kcj} \cup Z_{kjc} , \quad \forall c \in C - \{j, k\}. \quad (6.11)$$

We may be interested in particular, when the number of centres is not too large, in the elementary cells that are drawn on the plane by the set of all hypercircles (usually lines) $Z_{cc'} \cap Z_{c'c}$, for any $c, c' \in C$. Those cells are like puzzle pieces, the reunion of some of which forms Z_k and the other choice areas. Each cell is characterized by a preference ordering of all the areas; cf. the work of Smith (1975) showing the equivalence between the GG model and a preferential structure.

6.2. Global comparison with other centres

Another possible definition of U_k is the following:

$$U_k = \sum (V_c : c \in C - \{k\}) . \quad (6.12)$$

That viewpoint is appropriate for distributive models; the amount of interaction with k , V_k , is here compared to the remaining interaction. When U_k is given by (6.12), the inequality $V_k \geq U_k$ expresses a property that is met in other fields [see e.g. Hanjoul and Thisse (1984)], namely that k exerts some

.6.8.

kind of dominance among the centres, or is in a state of absolute majority in terms of potential; whereas Z_k is the area where k is in a state of relative majority in terms of potential. We call the corresponding area the *dominance area* of k and denote it as Y_k .

There is a relation between Y_k and Z_k :

$$Y_k \subseteq Z_k \subseteq Y_k [a_k(\#C - 1)] . \quad (6.13)$$

In this writing, $Y_k[a']$ means the dominance area of centre k in a system $a' = (a'_j: j \in C)$ of attractivenesses, and $a_k(\eta)$ is equal to the current system of attractivenesses, a , with a_k replaced by ηa_k , for some real η .

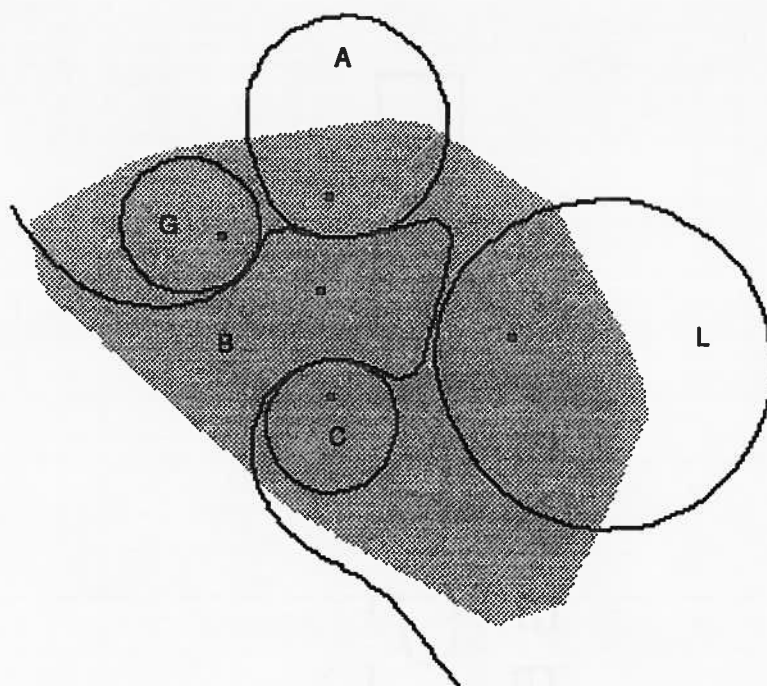


Fig. 6.5. — Dominance areas computed from Thill's model.

The first inclusion is obvious: in Y_k , we have:

$$a_k F_k \geq \sum (a_c F_c: c \in C - \{k\}) ; \quad (6.14)$$

hence, as the attractivenesses and the function F are positive numbers, we have $a_k F_k \geq a_c F_c$, $\forall c \in C$. As to the second: summing the latter inequalities, we find

$$(\#C - 1) a_k F_k \geq \sum (a_c F_c : c \in C - \{k\}) ;$$

hence the result.

In a similar perspective, U_k may be defined as the *arithmetic average* of the other potentials:

$$U_k = \sum (V_c : c \in C - \{k\}) / (\#C - 1) . \quad (6.15)$$

We call the corresponding area the *arithmetic average area* of k .

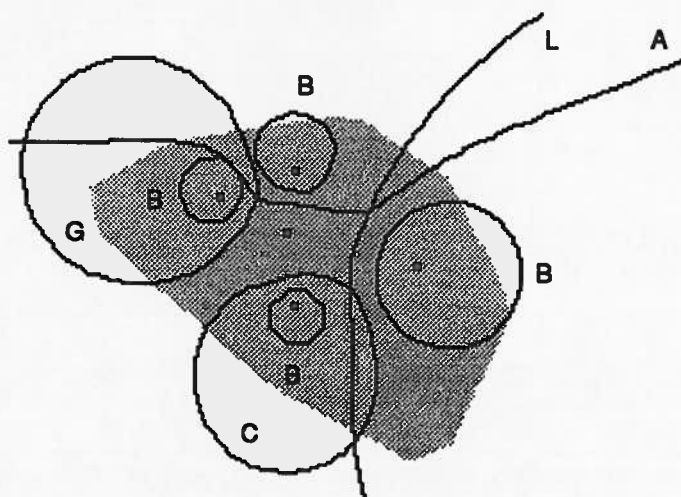


Fig. 6.6. — Arithmetic average areas computed from Thill's model.

The definitions of the dominance and arithmetic average areas are instances of a more general one:

$$V_k \geq \tau \sum (V_c : c \in C) , \quad (6.16)$$

for $\tau \in]0, 1[$. We denote the corresponding area as $Sh_k(\tau)$ and call it, as in Section 1.7, the τ -share area of k . The dominance area of k and its arithmetic

$$(1-\tau) a_k F_k \geq \tau \sum (a_c F_c: c \in C - \{k\}) ,$$

we have the equality

$$Sh_k(\tau) = Y_k [\underline{a}_k(1/\tau - 1)] \quad (6.17)$$

In particular, the arithmetic average area is equal to $Y_k[a_k(\#C - 1)]$; formula (6.13) thus expresses that the market area contains the dominance area but is contained in the arithmetic average area.

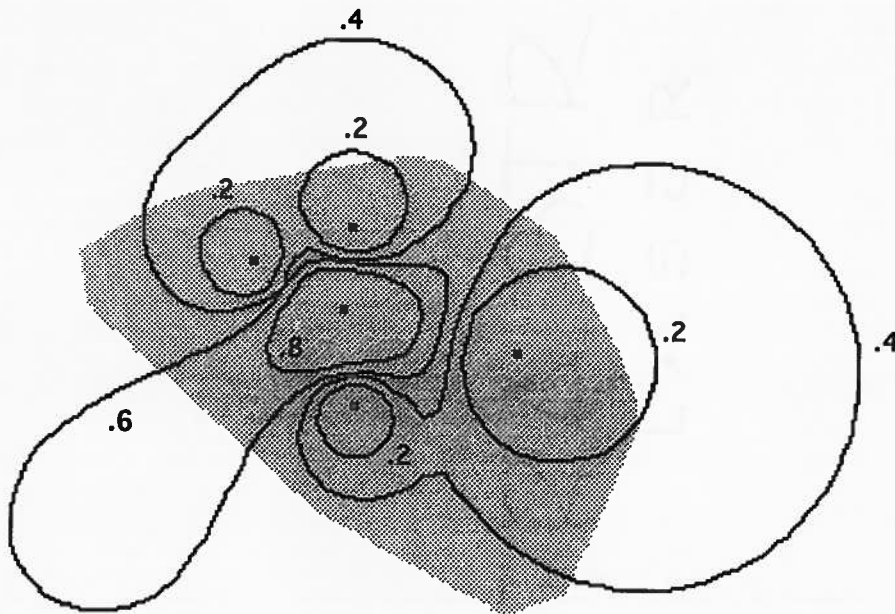


Fig. 6.7. — Share areas of Brussels at the levels 20%, 40%, 60%, and 80%, based on Thill's model.

The ratio $V_k/\Sigma (V_j: j \in C)$ is the *market share* of k ; which is also the *probability of interaction* with k in distributive models, as particularly noticed by Huff (op. cit.). As a matter of fact, that author, together with Wrigley (1977) and others, criticized the very idea of delimiting the area associated with a centre and considered that the set of all *isoshare* (or *isoprobability*) lines, which is equivalent to the set of all the τ -share areas for all possible values of τ , would be the only set really deserving the

name of market area, being a useful tool when assessing the total demanded addressed to a centre.

Because of (6.17), the properties of share areas may immediately be derived from those of dominance areas, as studied in Chapters 7 and 9. We also have this property:

(6.2) If $\tau \geq 50\%$, the τ -share areas $Sh_C(\tau)$ are isolated from one another;
if $\tau \leq 1/\#C$, their set is chained.

6.3. Other definitions

First, other types of averages than (6.8) and (6.15) may be thought of. Using a very common way to compute averages, we may define thus the area of k :

$$V_k \geq [\sum (V_C^\beta : c \in C) / \#C]^{1/\beta} \quad (\text{for some } \beta \neq 0). \quad (6.18)$$

(Notice that V_k may be compared, equivalently, with the average of the other potentials). Usual types of averages belong to that family: the *arithmetic* ($\beta = 1$), *quadratic* ($\beta = 2$), *harmonic* ($\beta = -1$), and even the *geometric* one. The latter, using which the area of k is defined by

$$V_k \geq [\prod (V_C : c \in C)]^{1/\#C}, \quad (6.19)$$

is indeed the limit of the general average when $\beta \rightarrow 0$. We shall refer to the corresponding areas as *average areas*¹. Of course, Prop. 6.2 may be used here; *mutatis mutandis*, it shows that the average areas are chained.

¹ Moreover, we may consider weighted averages. The arithmetic case, in particular, is described by $V_k \geq \sum (\omega_C V_C : c \in M) / \sum (\omega_C : c \in M)$ and is also, obviously, formally reducible to (6.16). Our statements should of course not be applied here if the weights depend on potentials. An exception occurs when they are defined thus: $\omega_C = 1/V_C$; we then get in fact a harmonic average.

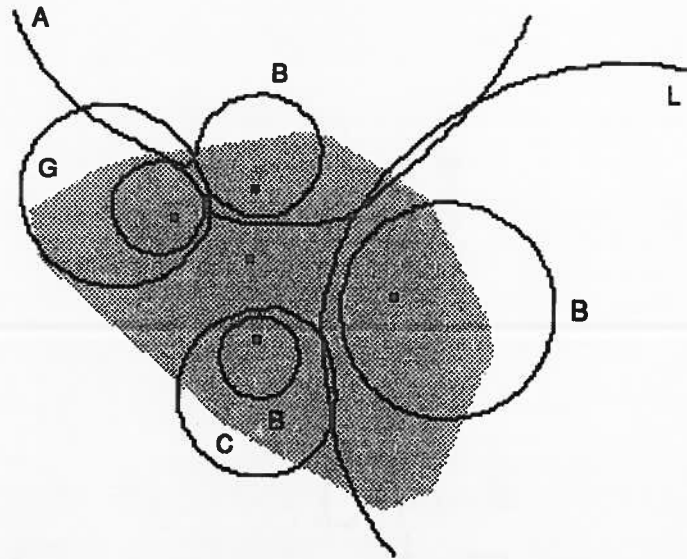


Fig. 6.8. — Quadratic average areas computed from Thill's model.

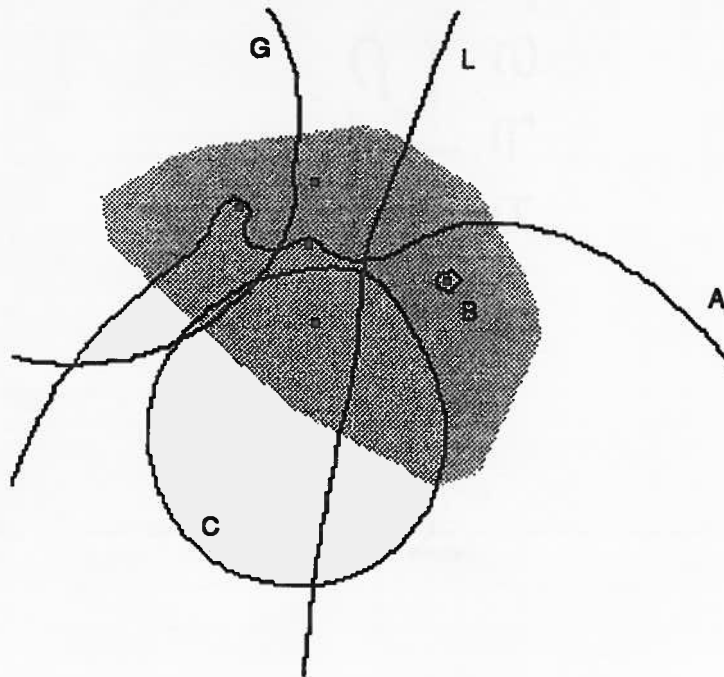


Fig. 6.9. — Geometric average areas computed from Thill's model.

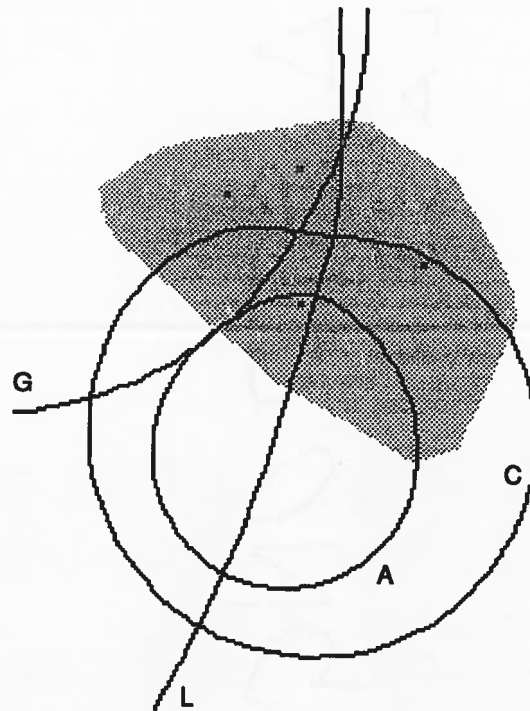


Fig. 6.10. — Harmonic average areas computed from Thill's model.
The area of Brussels is the whole plane.

Second, another group of definitions of the market area of a centre k are frequently met. The consumer locations, i , are ranked according to some index p_{ik} , and the market area of k is the smallest of the zero-dimensional areas $\{i \in M: p_{ik} \geq \gamma\}$ (for variable γ) that encompasses a given percentage ε of the total demand P_k , equal to $\sum(t_{i'k}: i' \in M)$, addressed to centre k ². It is usual, in this perspective, to distinguish a *primary* or *core* area (ε between 60 and 70%) and a *secondary* (ε between 15 and 25%, minus the primary area); the remaining part of M where the demand is nonzero is the *tertiary* or *fringe* area. We may study those areas in our perspective by replacing M by $\mathcal{R}^2$ in those definitions, dropping of course the requirement that $m_i \neq 0$. For us, the boundary of that kind of market area should thus be some isoline of p_{ik} .

Several definitions of p_{ik} are encountered in the literature. Goodchild (1992) uses as p_{ik} an assessment of consumer density in the vicinity

² Appelbaum (see below) called P_k the *drawing power* of k .

of the consumer location. He also suggests the use of the ratio t_{ik}/P_k . In the latter case, when all consumer locations have a same mass m , $\rho_{ik} = m a_k F_{ik}/P_k$; the market area is thus a disk. But, needless to say, our study can generally not be applied to areas so defined.

For Huff and Batsell (1977) and Huff and Rust (1984), on the contrary, ρ_{ik} is simply the distance δ_{ik} , so that the market areas should be disks anyway. Those authors, however, preferred to draw a complex curve closely encompassing the retained points of M .

For Appelbaum (1966), ρ_{ik} should be equal to the per capita sales, t_{ik}/Π_i , Π_i being the population of the site. In the GG model, we have

$$t_{ik} / \Pi_i = (m_i / \Pi_i) a_k F_k.$$

If we assume, as is frequently done, that m_i is proportional to Π_i , t_{ik}/Π_i is proportional to $a_k F_k$ and thus a decreasing function of δ_k . In that case, we conclude that the isolines of ρ_{ik} should simply be circles centred on k ; as they should in the work of Huff and Batsell, and Huff and Rust.

Things are different if we use a consumption-constrained model. Here, it is T_i , the total interaction in i , that is known a priori and supposed not to vary, and is likely to be proportional to Π_i . Thus t_{ik}/Π_i should be proportional to t_{ik}/T_i ; i.e., to $t_{ik}/\sum(t_{ic}: c \in C)$, the market share of k . In this case only we do find the correspondence mentioned by O'Kelly and Miller (1989): the isolines of t_{ik}/Π_i are *isoshare lines*. Appelbaum also drew isoshare lines for the example he presented. Noteworthily, they do not correspond at all with his isolines of t_{ik}/Π_i .

7. Starshapedness and position

7.1. Starshapedness

The *Voronoi cell* (also known as Thiessen polygon) of a centre c is so defined:

$$\mathcal{V}_k = \{i: \forall c \in C: \delta_k \leq \delta_c\} .$$

We also define the *anti-Voronoi cell* of k :

$$a\mathcal{V}_k = \{i: \forall c \in C: \delta_k \geq \delta_c\} .$$

Every point of the plane belongs to some Voronoi cell and to some anti-Voronoi.

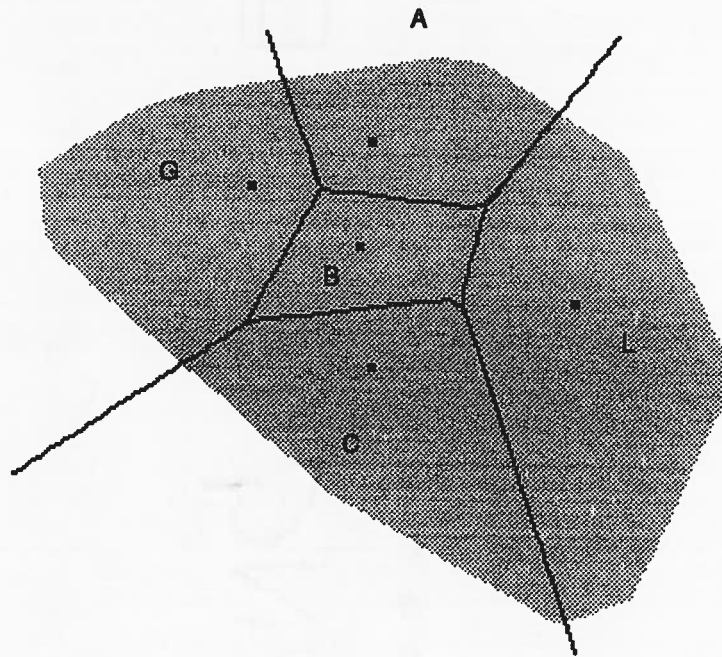


Fig. 7.1. — Voronoi cells of the cities of the third level of urban hierarchy in Belgium.

We also extend in several ways the notion of starshapedness:

- (1°) a set E is *starshaped* in a set F w.r.t. a point p iff. $\forall i \in E$:
 $[ip] \cap F \subseteq E$;
- (2°) E is *antistarshaped* w.r.t. k iff. $\forall i \in E$: the straight half-line
 $[ki) - [ki[\subseteq E$;

- (3°) E is *antistarshaped* w.r.t. k in a set F iff. $\forall i \in E: ([ki) - [ki[) \cap F \subseteq E$;
 (4°) E is *weakly starshaped* w.r.t. a set F iff. $\forall i \in E: \exists i' \in F: [ii'] \subseteq E$.

Notice that the antistarshapedness of E w.r.t. k means that E^- is starshaped w.r.t. k .

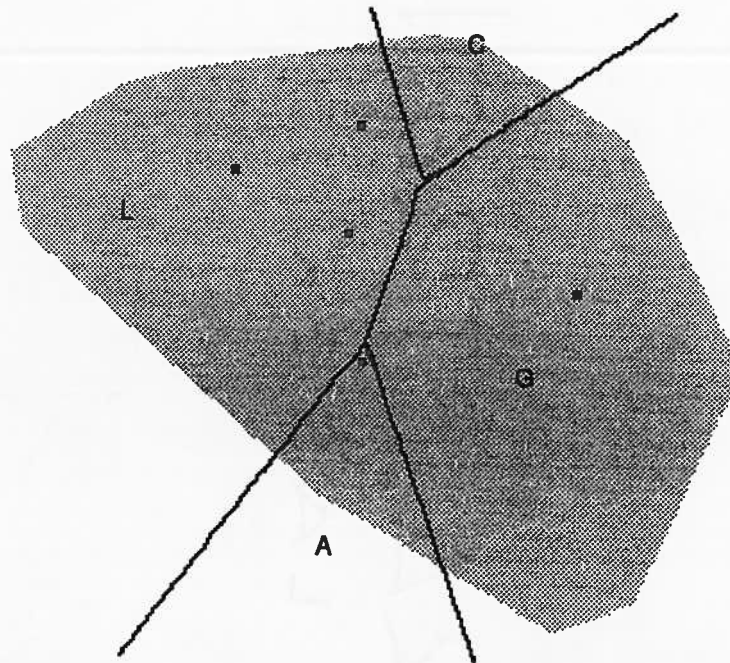


Fig. 7.2. — Anti-Voronoi celles of the cities of the third level of urban hierarchy in Belgium.

Last, we denote by m_{ck} the middle of $[ck]$. We then get what follows:

- (7.1) (a) The areas $Z_k \cap \bigcap_{c \in C-\{k\}} D_{m_{ck}}(\delta_{ck}/2)$
 and $Y_k \cap \bigcap_{c \in C-\{k\}} D_{m_{ck}}(\delta_{ck}/2)$ are starshaped w.r.t. k ;
 (b) When h is concave:
 (i) $Z_k \cap \mathcal{V}_k$ and $Y_k \cap \mathcal{V}_k$ are starshaped w.r.t. k ;
 (ii) if k is least attractive, Z_k is starshaped w.r.t. k ;
 (c) When h is convex:
 (i) Z_k and Y_k are starshaped w.r.t. k in $\mathcal{V}_k$;
 (ii) if k is most attractive, Z_k is starshaped w.r.t. k ;

(d) When h is linear: Z_k and Y_k are starshaped w.r.t. k , and Z_k^- is weakly starshaped w.r.t. $C - \{k\}$.

Let us first consider the proof concerning Z_k . Item (a) is the direct extension of Prop. 5.4c. The proof of (d) is immediate: when h is linear, it is concave and convex at the same time, so that $k \in Z_{kc}$ whether k is more attractive than c or not; the rest follows from Prop. 3.2. The items (b) and (c) are direct consequences of Prop. 5.4.

Now concerning Y_k , the inequality defining it, according to (6.14), may be written:

$$a_k \geq \sum_{c \in C - \{k\}} a_c F_c / F_k;$$

i.e.,

$$a_k \geq \sum_{c \in C - \{k\}} a_c / \exp (h_c - h_k) .$$

We may thus study the variation of the differences $h_c - h_k$ for a radial move from k . Instead of Prop. 5.4, we thus use its demonstration. Apart from this, the proof is obvious. Q.E.D.

Clearly, all those properties are valid whatever the transportation space under consideration. The sole restriction, in Prop. 7.1a, is that $D_{m_{ck}}(\delta_{ck}/2)$ has to be replaced, more generally, by the set where any move toward k decreases δ_c and vice versa, any move toward c decreases δ_k .

Another remark: the area $\bigcap \{D_{m_{ck}}(\delta_{ck}/2) : c \in C - \{k\}\}$ which appears in item (a) always contains $[kk']$, in which k' denotes the point closest to k amongst the convex hull of $C - \{k\}$ (i.e., the smallest convex area containing that set). If k belongs to that convex hull (i.e., if k is not a corner of the convex hull of C), that intersection is reduced to $\{k\}$ and item (a) is useless; see Section 9.2.

Denote now by T_{kc} the half plane containing k and the boundary of which is perpendicular to (ck) in c ; by σT_{kc} , the other half plane having the same boundary; and let

$$\mathcal{W}_k = \bigcap_{c \in C - \{k\}} \mathcal{T}_{kc} ,$$

$$a\mathcal{W}_k = \bigcap_{c \in C - \{k\}} a\mathcal{T}_{kc} ,$$

$$\mathcal{U}_k = \bigcap_{c \in C - \{k\}} \mathcal{T}_{ck} ,$$

and

$$a\mathcal{U}_k = \bigcap_{c \in C - \{k\}} a\mathcal{T}_{ck} .$$

Being intersections of half planes, all those areas are convex. They have the following properties, the proof of which is easy and left to the reader.

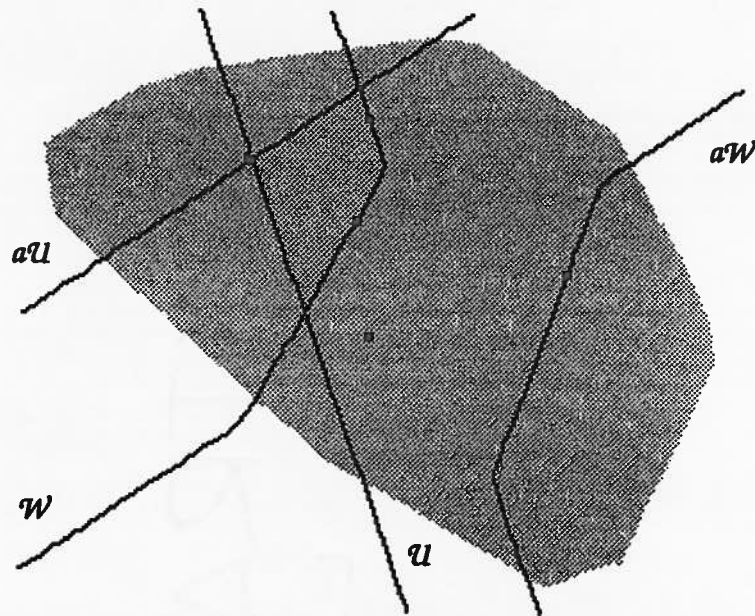


Fig. 7.3. — Areas $\mathcal{W}$, $a\mathcal{W}$, $\mathcal{U}$, and $a\mathcal{U}$:

the case of Gent in the third level of the Belgian urban hierarchy.

$\mathcal{W}_k$ is obviously homothetic to the Voronoi cell of k , $\mathcal{V}_k$, w.r.t. k and with a ratio 2; so that the area of $\mathcal{W}_k$ is four times that of $\mathcal{V}_k$, if those sets are bounded. The areas of this type are thus chained. $a\mathcal{W}_k$ is similarly homothetic, w.r.t. k , to the anti-Voronoi cell of k , $a\mathcal{V}_k$, so that $a\mathcal{W}_k$ is more remote from k than $a\mathcal{V}_k$. The effect is that the areas $a\mathcal{W}_k$ are like isolated infinite continents (or empty). Moreover, the convex hull of C (which is, as we said above, the smallest convex area containing C) lies between k and $a\mathcal{W}_k$, with which it has no intersection point, except maybe at the point of C most

distant from k ; C is indeed contained in T_{kc} , c being a centre most remote from k , whereas $a\mathcal{W}_k \subset aT_{kc}$.

Similarly, the intersection of the area $a\mathcal{U}_k$ with the convex hull of C is equal to $\{k\}$. The areas $\mathcal{U}_k$ and $a\mathcal{U}_k$ are sectors opposed by their summit, k . It is easily seen that they are reduced to $\{k\}$ if k belongs to the interior of the convex hull of C . If k belongs to the boundary of the convex hull but is none of its corners, $\mathcal{U}_k$ and $a\mathcal{U}_k$ are half-lines perpendicular to that boundary. When k is a corner of the convex hull, $\mathcal{U}_k$ and $a\mathcal{U}_k$ have a nonempty interior; the two straight lines forming their boundaries are perpendicular to the boundary of the convex hull.

The next lemma considers the concavity or convexity of h and of two derived functions, $h \circ \sqrt{\cdot}$ and $h \circ \exp$. Remember that the following implications obtain:

$$\begin{aligned} h \circ \sqrt{\cdot} \text{ is convex} &\Rightarrow h \text{ is convex} \Rightarrow h \circ \exp \text{ is convex} \\ h \circ \sqrt{\cdot} \text{ is concave} &\Leftarrow h \text{ is concave} \Leftarrow h \circ \exp \text{ is concave} . \end{aligned}$$

Of these properties, only the four combinations appearing in this lemma are consequently possible.

LEMMA 7.1. Consider two centres j and k ; $\Delta h = h_j - h_k$. Along any line passing through k :

- (a) When $h \circ \sqrt{\cdot}$ is derivable and convex,
 $d\Delta h/d\delta_k \geq 0$ in aT_{jk} ;
- (b) When $h \circ \sqrt{\cdot}$ is derivable and concave but h is convex,
 $d\Delta h/d\delta_k \leq 0$ in T_{jk} ;
- (c) When h is concave derivable and but $h \circ \exp$ is convex,
 $d\Delta h/d\delta_k \leq 0$ in T_{kj} ;
- (d) When $h \circ \exp$ is derivable and concave,
 $d\Delta h/d\delta_k \geq 0$ in aT_{kj} .

We know indeed that

.7.6.

$$d\Delta h/d\delta_k = h'_j d\delta_j/d\delta_k - h'_k = h'_j \cos(j\hat{i}k) - h'_k .$$

We thus see that $d\Delta h/d\delta_k \geq 0$ [resp. ≤ 0] iff.

$$h'_j \cos(j\hat{i}k) \geq h'_k \quad [\text{resp. } \leq h'_k]. \quad (7.1)$$

Let us begin with item (a). We have

$$(h \circ \sqrt{\cdot})'(\delta^2) = h'(\delta) / 2\delta ,$$

so that (7.1) may be written:

$$(h \circ \sqrt{\cdot})'_j \delta_j \cos(j\hat{i}k) \geq (h \circ \sqrt{\cdot})'_k \delta_k . \quad (7.2)$$

On $\mathcal{AT}_{jk}$, $\delta_j \geq \delta_k$. If $h \circ \sqrt{\cdot}$ is convex, we thus have, in this area, $(h \circ \sqrt{\cdot})'_j \geq (h \circ \sqrt{\cdot})'_k$. Consequently, as $\cos(j\hat{i}k) \geq 0$ on $\mathcal{AT}_{jk}$, the inequality $\delta_j \cos(j\hat{i}k) \geq \delta_k$, which defines $\mathcal{AT}_{jk}$, implies (7.2).

Let us now come to item (b). Suppose first that we are in $D_0(\delta_{ok})$, which is contained in $\mathcal{T}_{jk}$. Here, $\cos(j\hat{i}k) \leq 0$, which entails that $d\Delta h/d\delta_k \leq 0$. On the other hand, in the region defined by $\delta_j \leq \delta_k$, we have $h'_j \leq h'_k$, since h is convex; thus, according to (7.1), $d\Delta h/d\delta_k \leq 0$ in that half-plane, which is also a subset of $\mathcal{T}_{jk}$. In the remaining part of $\mathcal{T}_{jk}$, we have $(h \circ \sqrt{\cdot})'_j \leq (h \circ \sqrt{\cdot})'_k$ since $h \circ \sqrt{\cdot}$ is concave and $\delta_j \geq \delta_k$. We also have $\cos(j\hat{i}k) \geq 0$ and, according to the definition of $\mathcal{T}_{jk}$, $\delta_j \cos(j\hat{i}k) \leq \delta_k$. The opposite of (7.2) thus holds, and $d\Delta h/d\delta_k \leq 0$. The item is thus completely proved.

Item (c) is also immediately proved concerning $D_0(\delta_{ok})$ and, since h is now concave, the subset of $\mathcal{T}_{kj}$ where $\delta_j \geq \delta_k$. As to the remaining area, notice first that

$$(h \circ \exp)'[\ln(\delta)] = \delta h'(\delta) ,$$

so that (7.1) may also be written

$$(h \circ \exp)'_j \delta_k \cos(j\hat{i}k) \geq (h \circ \exp)'_k \delta_j . \quad (7.3)$$

The convexity of $h \circ \exp$ implies that $(h \circ \exp)'_j \leq (h \circ \exp)'_k$ on the area under consideration, since $\delta_j \leq \delta_k$. But, as we are in $\mathcal{T}_{kj}$, we also have

$\delta_k \cos(j\hat{i}k) \leq \delta_j$, and $\cos(j\hat{i}k)$ is here ≥ 0 ; the opposite of (7.3) thus holds, and $d\Delta h/d\delta_k \leq 0$.

Last, on $\mathcal{AT}_{kj}$, the region considered in item (d), $\delta_j \leq \delta_k$; $h \circ \exp$ being concave, we thus have $(h \circ \exp)'_j \geq (h \circ \exp)'_k$. Since that area is defined by $\delta_k \cos(j\hat{i}k) \geq \delta_j$ with $\cos(j\hat{i}k) \geq 0$, (7.3) obtains in $\mathcal{AT}_{kj}$. This completes the proof of the whole lemma. Q.E.D.

Most interesting are the intermediate cases; between (a) and (b): $h(\delta) = \delta^2$, between (b) and (c): $h(\delta) = \delta$, and between (c) and (d): $h(\delta) = \ln(\delta)$. The lemma does in fact not bring us any new information about the first two functions. But it does in the case of $\ln(\delta)$ when k is the more attractive centre (Fig. 7.4). On a purely geometric level, it implies that any Apollonian disk relative to j and k and containing j cuts the perpendicular to (jk) at j in two points that are the points of contact of the tangents issued from k to the disk; the disk is indeed visible from k up to the straight line determined by those contact points, but not beyond. Otherwise said, if we consider a disk, some point k outside it, and j , the intersection point of the normal issued from k to the disk and of the polar of k w.r.t. the disk, then the disk is Apollonian w.r.t. j and k . Or, expressed in the most simple way: *two points are Apollonian foci of a disk iff. they are conjugated w.r.t. the disk and aligned with its centre.*

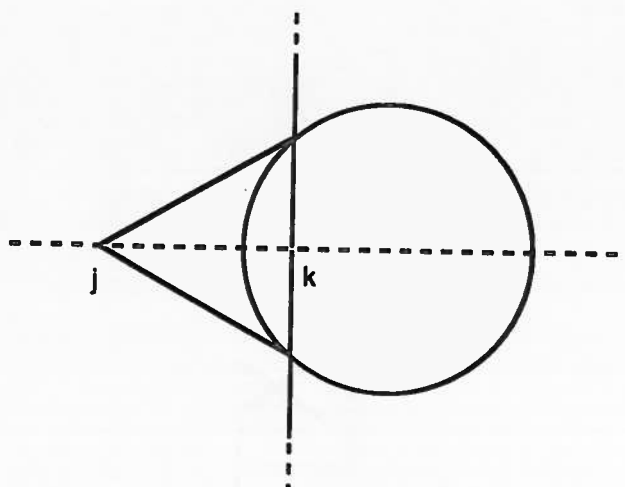


Fig. 7.4. — Application of Lemma 7.1 when h is logarithmic.

Notice also that the transportation cost functions of the type δ^a correspond to the first item when $a \geq 2$; to item (b), when $1 \leq a \leq 2$; and to item (c), when $0 < a \leq 1$. The function $-\delta^{-a}$, with $a > 0$, corresponds to item (d).

Of course, the properties listed in Lemma 7.1 could have been introduced in Part I, but their interest would not have been as obvious as it is now. On its basis, we may indeed construct this new theorem:

- (7.2) (a) If $h \circ \sqrt{\cdot}$ is derivable and convex,
 $Z_k \cap \mathcal{U}_k$ and $Y_k \cap \mathcal{U}_k$ are antistarshaped w.r.t. k ;
- (b) If h is derivable and convex but $h \circ \sqrt{\cdot}$ is concave,
 $Z_k \cap \mathcal{U}_k$ and $Y_k \cap \mathcal{U}_k$ are starshaped w.r.t. k ;
- (c) If h is derivable and concave but $h \circ \exp$ is convex,
 $Z_k \cap \mathcal{W}_k$ and $Y_k \cap \mathcal{W}_k$ are starshaped w.r.t. k ;
- (d) If $h \circ \exp$ is derivable and concave,
 $Z_k \cap \mathcal{W}_k$ and $Y_k \cap \mathcal{W}_k$ are antistarshaped w.r.t. k .

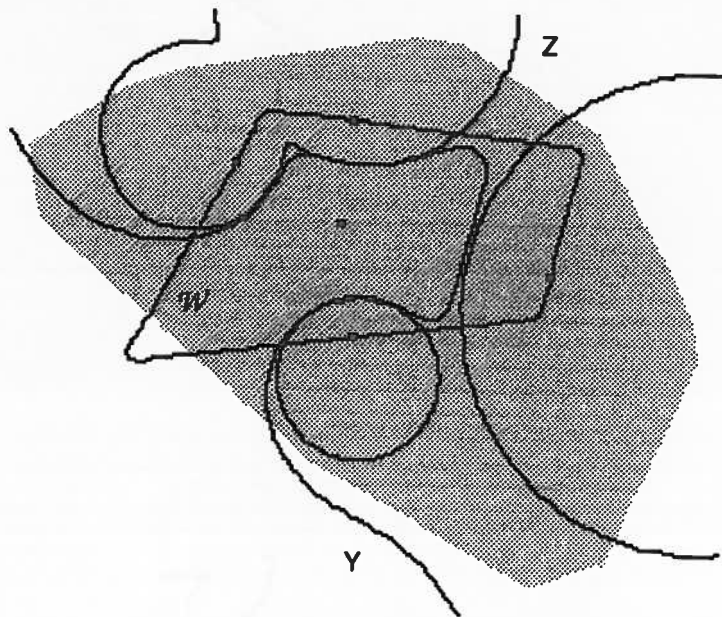


Fig. 7.5. — Market and dominance areas (computed on basis of Thills model) and area $\mathcal{W}$ for Brussels in the third level of urban hierarchy in Belgium.

Interestingly, with t.c.f.'s concerned by Prop. 7.2c, it is easily shown that all market areas are starshaped by their centres when those are organized in some regular patterns, like that described in Fig. 7.6. In that example, we face an infinite regular hexagonal pattern of centres; market areas are determined according to a Pareto model: $F(\delta) = 1/\sqrt{\delta}$, and there are two levels of attractiveness: 1.3 (level I) and 2 (level II). The market areas of level I are starshaped according to Prop. 7.1b (and even convex according to Prop. 8.1b). Those of level II are also starshaped by their centres. The market area of such a centre, say k , is indeed contained in the hexagon that would constitute its Voronoi cell if we only considered the level-II centres; as that hexagon is contained in the hexagon $\mathcal{W}_k$, we have $Z_k \subseteq \mathcal{W}_k$, and the property results from Prop. 7.2c. This may be of course be applied to central place models like those studied by Parr (1995b).

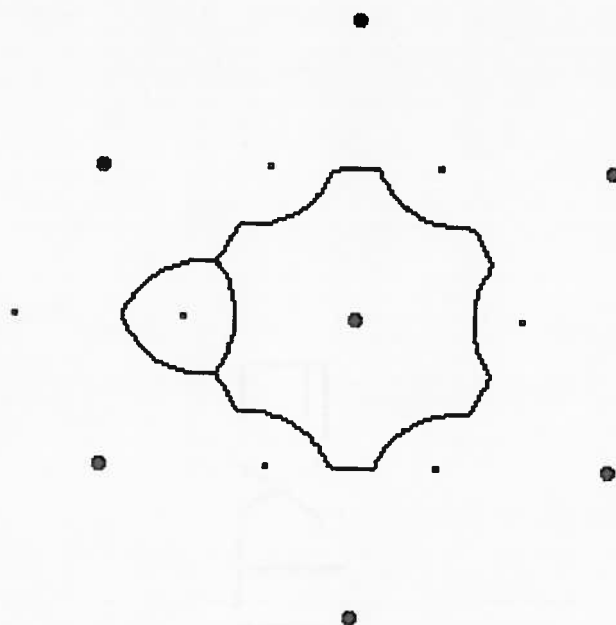


Fig. 7.6. — Two market areas in a Pareto model
with an infinite regular hexagonal pattern of centres and two levels of attractiveness.

When we look at already drawn market areas and want to check if one of them, Z_k , has the properties of starshapedness predicted by the theory, it is not always necessary to take all the other centers into account. We have already seen in Chapter 6 that, if we restrict the study to G , the set of the centres having a market area with nonempty interior, we may replace C by G in all the statements concerning centres $\in G$. Let us now restrict G to two smaller sets:

$$G_k = \{c \in G: \delta(Z_k, Z_c) = 0\}$$

(7.4)

$$G_k^{(0)} = \{c \in G: \delta(Z_k^0, Z_c) = 0\}$$

in which definition we denote by $\delta(E, F)$ the distance between the sets E and F ; i.e., the infimum of the set of the distances between points of E and points of F . Otherwise said, concerning G_k : it is the set of the centres corresponding to market areas adjacent to Z_k .

(7.3) If $k \in G$ and if $Z_k \langle G_k \rangle$ [resp. $Z_k^0 \langle G_k^{(0)} \rangle$] is starshaped w.r.t. k , then $Z_k = Z_k \langle G_k \rangle$ [resp. $Z_k^0 = Z_k^0 \langle G_k^{(0)} \rangle$].

Restricting G to G_k may let appear, when we compute $Z_k \langle G_k \rangle$, new points that were initially not in Z_k ; we have thus $Z_k \subseteq Z_k \langle G_k \rangle$. But those points can appear only in $\cup \{Z_c: c \in G - G_k\}$. Consequently, $\delta(Z_k, Z_k \langle G_k \rangle - Z_k) \geq \min\{\delta(Z_k, Z_c): c \in G - G_k\}$, which is > 0 according to the definition of G_k . This is even true, *lato sensu*, when $Z_k = Z_k \langle G_k \rangle$, since we have then $\delta(Z_k, Z_k \langle G_k \rangle - Z_k) = \delta(Z_k, \emptyset) = \min \emptyset = \infty$.

Suppose now that there is some $i \in Z_k \langle G_k \rangle - Z_k$ and that $Z_k \langle G_k \rangle$ is starshaped w.r.t. k . Hence, $[ik] \subseteq Z_k \langle G_k \rangle$; every point of $[ik]$ is thus either in Z_k or in $Z_k \langle G_k \rangle - Z_k$. This implies, as $i \notin Z_k$ and $k \in Z_k$, that $\delta(Z_k, Z_k \langle G_k \rangle - Z_k) = 0$, in contradiction with what we have established above. $Z_k \langle G_k \rangle - Z_k$ is thus empty; i.e., $Z_k = Z_k \langle G_k \rangle$.

Concerning $Z_k \langle G_k^{(0)} \rangle$, we just replace Z_k by Z_k^0 and $Z_k \langle G_k \rangle$ by $Z_k^0 \langle G_k^{(0)} \rangle$, and the proof remains valid. **Q.E.D.**

As a consequence, if we consider only the centres contained in G_k and so deduce that $Z_k \langle G_k \rangle$ must theoretically be starshaped w.r.t. k , we may apply that conclusion to Z_k itself.

7.2. Position

We define that k is a *dominant centre* iff.

$$a_k \geq \sum_{c \in C - \{k\}} a_c ;$$

k is *strictly dominant* iff. the inequality is strict. This is the vocabulary of the majority theorem of firm location theory; see Hanjoul and Thisse (1984). Of course, supposing there are at least three centres,

$$\begin{aligned} k \text{ is dominant} & \Rightarrow k \text{ is most attractive} \\ k \text{ is least attractive} & \Rightarrow k \text{ is not dominant.} \end{aligned}$$

Notice also that there is always at least one most attractive centre and one least attractive, but that there is one dominant centre at most (possibly no dominant centre at all). In Belgium, in the study of Thill, Brussels is the dominant centre with 53.8% of the total attractiveness of the cities of the third level of the urban hierarchy ¹. We have the following properties:

- (7.4) (a) If k is least attractive, $Z_k \subseteq \mathcal{V}_k$;
 (b) If k is not dominant, $Y_k \subseteq a\mathcal{V}_k^-$;
 (c) If k is most attractive, $\mathcal{V}_k \subseteq Z_k$;
 (d) If k is dominant, $\mathcal{V}_k \subseteq Y_k$ ².

Items (a) and (c) are basic consequences of our study of the two-center case, namely of Prop. 2.3. The proof of (b) and (d) is obvious; let us prove (b) for instance. Take any point i of $a\mathcal{V}_k^-$; thus, $\forall c \in C: \delta_k \geq \delta_c$, and $F_k \leq F_c$. The inequality $a_k < \sum_{c \in C - \{k\}} a_c$ then implies that $a_k F_k < \sum_{c \in C - \{k\}} a_c F_c$; so that the point $i \notin Y_k$. This proves (b) by contraposition. **Q.E.D.**

Concerning the belonging of a centre to its own market area, we have:

¹ We deduce that percentage and the others from his theoretical areas.

² This is in fact a particular case of an obvious property: if $a_k / \sum_{c \in C} a_c \geq \varepsilon / (1 + \varepsilon)$, then the area $\{i \in \mathbb{R}^2: \forall c \in C - \{k\}: F_c / F_k \leq \varepsilon\} \subseteq Y_k$.

- (7.5) (a) (i) If $h(0) = -\infty$, $k \in Y_k$;
- (ii) if k is most attractive, $k \in Z_k^0$;
- (iii) if k is dominant, $k \in Y_k^0$;
- (b) When h is concave:
- (i) if $Z_k \neq \emptyset$, $k \in Z_k$;
- (ii) if $Y_k \cap \mathcal{V}_k \neq \emptyset$, $k \in Y_k$;
- (iii) if $h \circ \exp$ is convex and $Y_k \cap \mathcal{W}_k \neq \emptyset$, $k \in Y_k$;
- (c) When h is convex and $h \circ \sqrt{\cdot}$ is concave:
- (i) if $Z_k \cap \mathcal{U}_k \neq \emptyset$, $k \in Z_k$;
- (ii) if $Y_k \cap \mathcal{U}_k \neq \emptyset$, $k \in Y_k$;
- (d) When h is linear and $Y_k \neq \emptyset$, $k \in Y_k$.

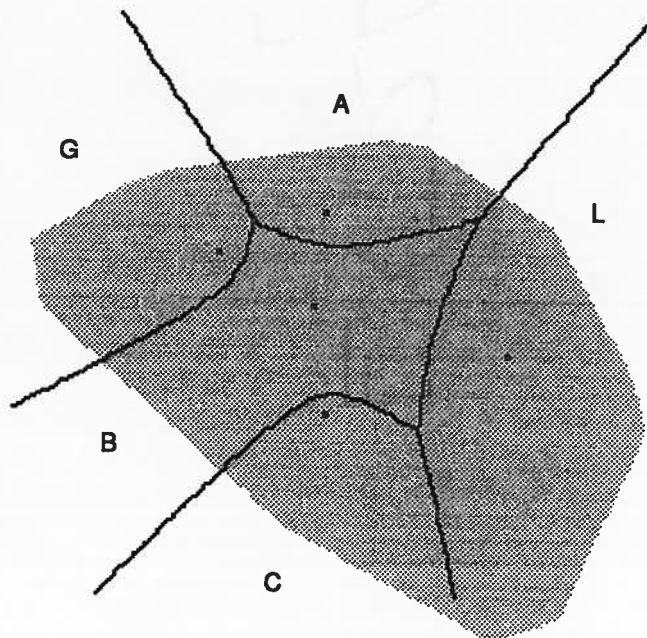
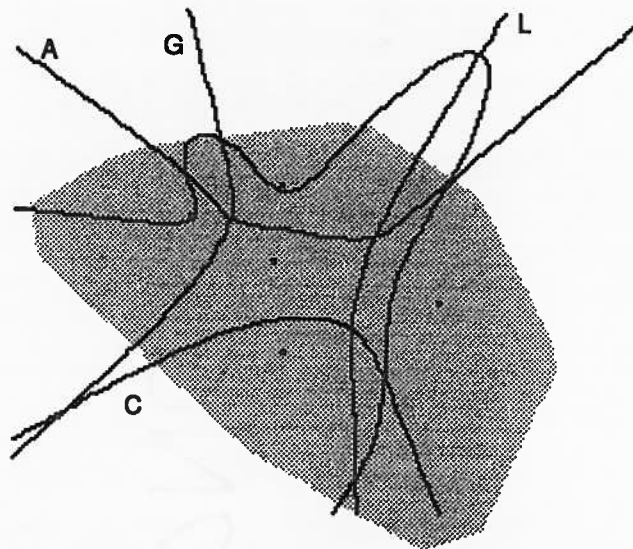


Fig. 7.7. — Market areas of the third level of urban hierarchy in Belgium when h is linear (exponential gravity model). The model was calibrated in order to maintain the breaking point of Thill's model between Brussels and Liège.

Item (a.i) is perfectly obvious. As to item (a.ii): we know that, when there are only two centres, the more attractive one (j in Part I) is contained in its market area; this is also true when both centres are equally attractive. That is why Z_{kc} contains k for every $c \in C - \{k\}$; consequently, (6.7) implies that $k \in Z_k$. Item (b.i) is similarly proved on the basis of Prop. 4.3b: in this case k belongs to Z_{kc} (supposed $\neq \emptyset$) even if k is less attractive than c . Items (a.iii), (b.ii), and (d) are direct consequences of Prop. 7.4d, 7.1b.i, and 7.1d, respectively. The remaining items: (b.iii) and (c), are also immediate aftermaths of Prop. 7.2c and 7.2b. **Q.E.D.**



7.8. — Arithmetic average areas; see Fig. 7.7.

Prop. 7.5 is illustrated in Fig. 7.7 and 7.8. On the other hand, the fact that Z_k may not contain its centre when h is not concave and when k is more attractive than some other centre belonging to C does not appear in that figure, but is well known; see Boots (1980).

7.3. About some finite-range distance deterrence functions

A natural extension of our usual assumptions about the distance deterrence function F is to allow it to have a finite range (also termed *refusal distance*, 'distance de refus'); i.e., a threshold value, δ , at and beyond which F is reduced to zero while being nonzero before. Otherwise said, the t.c.f. h

becomes ∞ at δ and remains beyond. That idea has already been used in the field of geography, at least by Christaller (1933) and Beckmann (1976b); it is implicit in the model of Isard (1960) presented in our Section 1.5. It is convenient, of course, to rewrite as follows the definitions of the market and dominance areas:

$$Z_k = \{i: \forall c \in C: V_k \geq V_c\} \cap D_k(\delta)$$

$$Y_k = \{i: V_k \geq \sum_{c \in C - \{k\}} V_c\} \cap D_k(\delta),$$

to avoid including in them the area where all influence vanishes. Notice that the redefinition (also valid in the case we have studied so far, in which $\delta = \infty$) does not affect the convexity nor the starshapedness w.r.t. k of these areas, since $D_k(\delta)$ itself is convex and starshaped w.r.t. k .

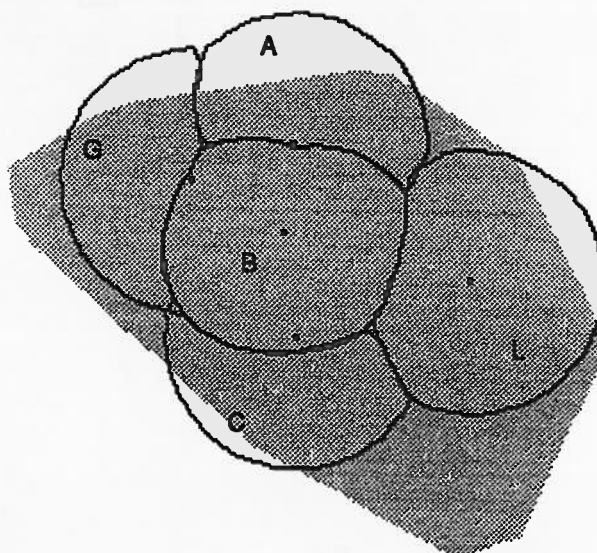


Fig. 7.9. — Theoretical market areas when $F(\delta) = \max \{0, \delta - \delta\}$ at the third level of urban hierarchy in Belgium. The parameter δ was computed so as to maintain the breaking point between Brussels and Liège in Thill's model. The t.c.f. is convex, so that Brussels starshapes its market area. The market areas obtained when $F(\delta) = \max \{0, \delta^2 - \delta^2\}$, with δ' computed in the same way, are very similar.

Two such deterrence functions are studied hereafter: $\max \{0, \delta - \delta\}$ and $\max \{0, \delta^2 - \delta^2\}$. It is easily seen that, in the former case, the t.c.f. is convex on $[0, \delta[$, whereas the associate function $h \cdot \sqrt{\cdot}$ is concave on $[0, \delta/2[$ and convex on $[\delta/2, \delta[$. In the second case, the t.c.f. is also convex on $[0, \delta[$,

and $h \cdot \sqrt{\cdot}$ is convex. The starshapedness of Z_k and Y_k may thus be partially studied through Prop. 5.4, 7.1 and 7.2. There is, however, a peculiar result concerning Y_k :

(7.6) When $F = \max \{0, \delta - (\cdot)\}$ or $F = \max \{0, \delta^2 - (\cdot)^2\}$: if k is dominant, Y_k is starshaped w.r.t. k .

The borders of the disks $D_c(\delta)$, for $c \in C$, divide $D_k(\delta)$ into a set of regions, $\mathfrak{S}_k$. To each region Ξ belonging to $\mathfrak{S}_k$ corresponds a set of centres, $C(\Xi)$, contained in C (and including k) and such that, for all point i of Ξ° and all $c \in C$, $V_c > 0 \Leftrightarrow c \in C(\Xi)$.

Consider first the case in which $F(\delta) = \max\{0, \delta - \delta\}$. Letting

$$U(\Xi) = a_k (\delta - \delta_k) - \sum_{c \in C(\Xi) - \{k\}} a_c (\delta - \delta_c),$$

so that Y_k is defined, inside Ξ , by the inequality $U(\Xi) \geq 0$, consider a point of $Y_k \cap \Xi$ and a move toward k , from i to $i' \in]ik] \cap Y_k \cap \Xi$. Studying the ensuing variations, we have

$$\begin{aligned} \Delta U(\Xi) &= -a_k \Delta \delta_k + \sum_{c \in C(\Xi) - \{k\}} a_c \Delta \delta_c \\ &\geq -a_k \Delta \delta_k - \sum_{c \in C(\Xi) - \{k\}} a_c |\Delta \delta_c|. \end{aligned}$$

We know, however, the triangular inequality: $|\delta_{ci'} - \delta_{ci}| \leq \delta_{ii'}$; i.e., $|\Delta \delta_c| \leq |\Delta \delta_k| = -\Delta \delta_k$. Consequently, if we define, for any set of centres E ,

$$s_k(E) = \sum_{c \in E - \{k\}} a_c, \quad (7.5)$$

we find

$$\Delta U(\Xi) \geq [a_k - s_k(C(\Xi))] (-\Delta \delta_k).$$

Since k is dominant, $a_k \geq s_k(C)$, and thus $a_k \geq s_k(C(\Xi))$; hence, $\Delta U(\Xi) \geq 0$. Hence, Y_k is starshaped w.r.t. k in Ξ . [Notice that the proof also holds when $C(\Xi) = \{k\}$, if we set that the sum of the empty sequence, $\sum \emptyset$, equals 0].

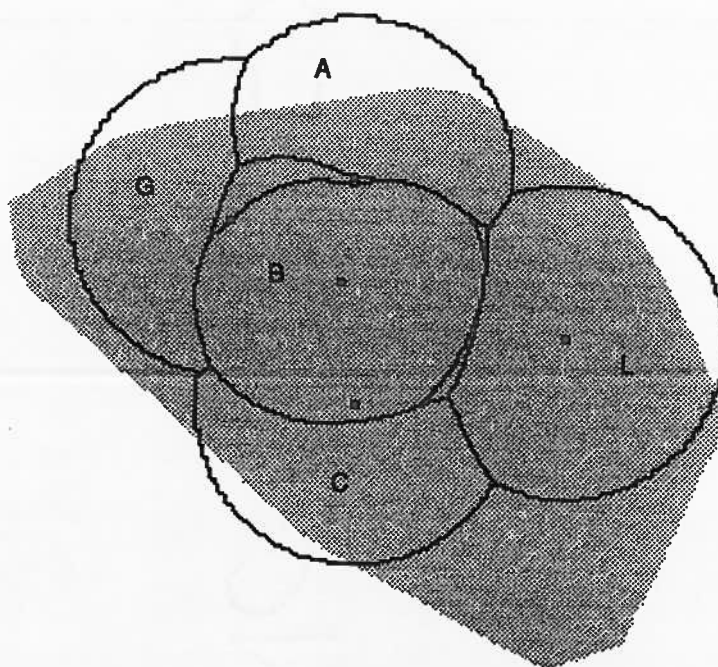


Fig. 7.10. — Theoretical dominance areas; see Fig. 7.9.

Consequently, if we move from i in the direction of k , we remain inside Y_k as long as we are inside Ξ ; i.e., until we reach either k (in which case our proof is achieved) or a point, say i' , of the boundary between Ξ and some other region $\Xi' \in \mathcal{S}_k$. As Y_k is similarly starshaped w.r.t. k in Ξ' , we may again progress radially toward k inside Y_k as long as we remain in Ξ' . And so on: since $[ik]$ has a finite number of intersection points with the boundaries of the disks $D_c(\delta)$, we must reach k after a finite number of similar steps, without having left Y_k , which is thus starshaped w.r.t. k .

Assume now that $F(\delta) = \max \{0, \delta^2 - \delta_k^2\}$. Inside the region $\Xi \in \mathcal{S}_k$, the inequality of Y_k writes:

$$a_k (\delta^2 - \delta_k^2) \geq \sum_{c \in C(\Xi) - \{k\}} a_c (\delta^2 - \delta_c^2);$$

i.e.,

$$a_k \delta_k^2 - \sum_{c \in C(\Xi) - \{k\}} a_c \delta_c^2 - \delta^2 (a_k - s_k[C(\Xi)]) \leq 0.$$

If we express the distances in terms of the coordinates x and y , it is clear that this inequality describes a half-plane when $a_k = s_k[C(\Xi)]$, or a disk when $a_k > s_k[C(\Xi)]$; there is no other possibility since k is dominant. In both cases, the inequality is that of a convex set. On the other hand, k belongs to

that area: this is immediately seen by setting $\delta_k = 0$ in the inequality (this is also implied by Prop. 7.5a.iii, since the sign of F bears no influence on its proof). The area is thus starshaped w.r.t. k ; hence, Y_k , identical with it inside Ξ , is starshaped w.r.t. k in Ξ . The remainder of the proof follows that concerning the t.c.f. $\max \{0, \underline{\delta} - \delta\}$. Q.E.D.

Notice that the proof concerning $\max \{0, \underline{\delta} - \delta\}$ does not depend on the particular transportation space we are in.

7.4. Choice, share, median, and average areas

Thanks to (6.9), the case of *choice areas* is easy: the whole Section 7.1, except for Proposition 7.3, remains valid if we replace Z_k by $Ch_k(m)$.

To study a τ -share area, we know from (6.17) that we have just to multiply a_k by $1/\tau - 1$ and apply sections 7.1 to 7.3. Nevertheless, it is interesting to see what the condition that k is dominant now means with respect to τ . That condition, $(1/\tau - 1) a_k \geq \sum (a_c: c \in C - \{k\})$, may be written as

$$\tau \leq a_k / \sum (a_c: c \in C) .$$

Otherwise said, the condition is now that the considered market share of k is not larger than k 's share in attractiveness. For the Belgian cities of the third level, according to Thill's study, those percentages approximately amount to 53.8% (Brussels), 19.0% (Antwerp), 13.3% (Liège), 7.6% (Gent), and 6.3% (Charleroi).

As to *average areas*, first, notice that the property mathematically equivalent to the dominance of k may be expressed as ' $a_k \geq \bar{a}$ ', where $\bar{a}$ is the corresponding average attractiveness. In our example, that threshold is of course 20% ($1/\#C$) for the arithmetic average; the condition is met only by Brussels, and nearly by Antwerp.

When $\beta > 0$, (6.18) is equivalent to an inequality of the type $V_k \geq U_k$ with U_k as in (6.15), the attractiveness of any centre c being a_c^β ; and the distance deterrence function, F^β . When $\beta < 0$, F^β is still positive, but

strictly decreasing; this seems likely to invert the properties. In fact, we may then transform (6.18) into:

$$-v_k^\beta \leq \sum (-v_c^\beta: c \in C) / \#C .$$

If there is some number K large enough for the function $F^\sim$, equal to $K - F^\beta$, to be > 0 in the region under study, that inequality may finally be rewritten as

$$a_k^\beta F^\sim_k \leq \sum (a_c^\beta F^\sim_c: c \in C) / \#C + K[a_k^\beta - \sum (a_c^\beta: c \in C - \{k\}) / \#C] . \quad (7.6)$$

It is thus now the area where $v_k \leq [\sum (v_c^\beta: j \in C) / \#C]^{1/\beta}$ which is described by (6.15) with attractivenesses a_c^β and distance deterrence function $F^\sim$, if we neglect the constant term in the right-hand member of (7.6). In the geometric case, taking the logarithm of both members of (6.19) leads us to the inequality

$$-h_k \geq \sum (-h_c: c \in C) / \#C + [p_k - \sum (p_c: c \in C) / \#C] .$$

Here again, if there is some number K large enough for the function F^* , equal to $K - h$, to be > 0 in the region under study, we rewrite the inequality:

$$F^*_k \geq \sum (F^*_c: c \in C) / \#C + [p_k - \sum (p_c: c \in C) / \#C] \quad (7.7)$$

Neglecting the constant term of the right-hand member, we see that the area where $v_k \leq [\prod (v_c: c \in C)]^{1/\#C}$ is described by (6.15) with F^* as distance deterrence function, all attractivenesses being equal to 1.

Except for 7.4a, 7.4d, 7.5a.iii, and 7.6, the proofs concerning Y_k use derivatives and are not affected by the presence of such constant terms. We may thus apply Prop. 7.1 and 7.2 to those areas. However, the former four properties, which are also those where the dominance of k is taken into account, must be reconsidered. First, there is the restriction, anyway, that Prop. 7.6 must be used only in the region $\cap \{D_c(\bar{Q}): c \in C\}$ since F^β now becomes infinite when F falls to zero. For the rest, when $\beta = 0$, in the geometric case, the four properties remain valid. When $\beta < 0$, it is easily seen that they are inverted: 'k is dominant' must be replaced by ' $a_k \leq \bar{a}$ ', etc.

The median area of k has different properties according to whether v_k is compared with an even or an odd number of centres. For instance, in Thill's study of the cities of the third level in Belgium, a city may be compared with

the median of the four other cities of the same level, or with the median of all the five cities of the level. In the latter case, as we have said in Section 6.2, the median area of k is $Ch_k[(n+1)/2]$ and the reader is referred to what we have said in the beginning of the present section. In the former, if we consider the variation of δ_k on a straight half-line originating in k , the problem is piecewise similar to that of the study of an arithmetic average area (since V_k is compared with the arithmetic average of two other centres) and thus of Y_k . The dominance of k , when required, is here to be replaced by the condition:

$$a_k \geq (a_c + a_{c'})/2$$

where c and c' are the two most attractive (in the common meaning of that expression) among the set of centres chosen for the comparison, either C or $C-\{k\}$. Of course, that condition is verified by the most attractive centre; but certainly not by the centres the attractivenesses of which are below the level of the two most attractive centres. In our example, when every city is compared with the four others, those threshold attractivenesses are approximately equal to 16.1% (Brussels), 33.5% (Antwerp), and 36.4% for each of the three remaining cities. The condition is thus met by Brussels only.

8. Convexity and arcwise-connectedness

We shall first discuss the properties of market areas, and come back to the other types of areas in Section 8.4.

8.1. Convexity

We first establish a few preliminary results. Let E be some subset of the plane.

LEMMA 8.1. *If E is closed and convex and if $E^\circ \neq \emptyset$, $E = \text{cl}(E^\circ)$.*

Being nonempty, E° contains an infinity of points; consider thus a point a of E , and a point b of $E^\circ - \{a\}$. For some $\varepsilon > 0$, $D_b(\varepsilon) \subseteq E^\circ$; and, since E is convex, $\cup\{[ai]: i \in D_a(\varepsilon)\} \subseteq E$. Point a is clearly adherent to the interior of that reunion and thus belongs to $\text{cl}(E^\circ)$. E is thus contained in $\text{cl}(E^\circ)$. On the other hand, the fact that $E^\circ \subseteq E$ entails that $\text{cl}(E^\circ) \subseteq \text{cl}(E) = E$. Consequently, $E = \text{cl}(E^\circ)$. **Q.E.D.**

LEMMA 8.2. *E° is convex $\Leftrightarrow \text{cl}(E^\circ)$ is convex.*

The lemma is obvious when E° is empty: in that case, $\text{cl}(E^\circ) = E^\circ = \emptyset$, which is convex. Suppose now that $\#E^\circ > 1$. E° having more than one point, so has $\text{cl}(E^\circ)$, since $E^\circ \subseteq \text{cl}(E^\circ)$.

($\Rightarrow$) Let thus a and b be two distinct points of $\text{cl}(E^\circ)$, and consider any number $\varepsilon \in]0, \delta_{ab}/2[$; so $D_a(\varepsilon) \cap D_b(\varepsilon) = \emptyset$. By definition of $\text{cl}(E^\circ)$, $D_a(\varepsilon)$ contains some point i , and $D_b(\varepsilon)$ some i' , with i and i' in $E^\circ - \{a, b\}$. As E° is convex, $[ii'] \subseteq E^\circ$. Take now any point i'' of $[ab]$; obviously, $D_{i''}(\varepsilon) \cap [ii'] \neq \emptyset$. It is thus clear that $i'' \in \text{cl}(E^\circ)$ and that $[ab] \subseteq \text{cl}(E^\circ)$, so that $\text{cl}(E^\circ)$ is convex ¹.

¹ Notice that this part of the property holds also when E is not open.

($\Leftarrow$) Consider any two distinct points, a and b , of E° . There is some $\varepsilon > 0$ for which $D_a(\varepsilon)$ and $D_b(\varepsilon)$ are contained in E° , and thus in $\text{cl}(E^\circ)$. This set being convex, $\cup\{[ii'] : i \in D_a(\varepsilon), i' \in D_b(\varepsilon)\}$ is contained in $\text{cl}(E^\circ)$; and it is clear that $[ab]$ is in the interior of that reunion, and so in the interior of $\text{cl}(E^\circ)$; i.e., in E° , which is thus convex. Q.E.D.

LEMMA 8.3. If E is closed and convex, $E = \text{cl}(E^\circ)$ and E° is convex.

This lemma is a direct consequence of the two preceding ones. Q.E.D.

Consider now these sets of centres:

$$G'_k = \{c \in G; \#(Z_k \cap Z_c) = \aleph_1\},$$

$$G'^{(o)}_k = \{c \in G; \#(\text{cl}(Z_k^\circ) \cap Z_c) = \aleph_1\},$$

where $\aleph_1$ ('aleph one') is the transfinite number $\aleph$. In particular, $\#(Z_k \cap Z_c) = \aleph_1$ [resp., $\#(\text{cl}(Z_k^\circ) \cap Z_c) = \aleph_1$] when $Z_k \cap Z_c$ [resp., $\text{cl}(Z_k^\circ) \cap Z_c$] contains a path. It is even probably the only case, but we have found no satisfying proof of this. Notice also that

$$G'_k \subseteq G_k,$$

$$G'^{(o)}_k \subseteq G_k^{(o)},$$

the sets G_k and $G_k^{(o)}$ having been defined in (7.4). Indeed, A and B being two sets, $\text{cl}(A) \cap \text{cl}(B) \neq \emptyset \Rightarrow \delta(A, B) = 0$ [but it is possible to find sets A and B for which $\delta(A, B) = 0$ but $\text{cl}(A) \cap \text{cl}(B) = \emptyset$].

- (8.1) (a) If $Z_k \langle G'_k \rangle$ [resp. $Z_k^\circ \langle G'^{(o)}_k \rangle$] is convex and $k \in G$, $Z_k = Z_k \langle G'_k \rangle$ [resp. $Z_k^\circ = Z_k^\circ \langle G'^{(o)}_k \rangle$];
- (b) When the t.c.f. is convexifying and $k \in G$: if $k \preceq c \forall c \in G'_k$ [resp. $G'^{(o)}_k$], Z_k [resp. Z_k°] is convex;
- (c) When the t.c.f. is strictly convexifying and $k \in G$: Z_k is convex (resp. strictly convex) iff. $k \preceq c$ (resp. $k \ll c$) $\forall c \in G'_k$;
- (d) When the t.c.f. is $(.)^2$, Z_k is convex, but not strictly.

As to the first item, we give here the proof concerning Z_k and $Z_{k\langle G'_k \rangle}$. That concerning Z_k° and $Z_{k^\circ\langle G'_k \rangle}$ is similar; we just have to replace Z_k by Z_k° and $Z_{k\langle G'_k \rangle}$ by $Z_{k^\circ\langle G'_k \rangle}$.

Since $G'_k \subseteq G$, $Z_k \subseteq Z_{k\langle G'_k \rangle}$. The points of $Z_{k\langle G'_k \rangle} - Z_k$, if any, necessarily appear in areas Z_c , for some centres $c \in G - G'_k$. Otherwise said, $Z_{k\langle G'_k \rangle} - Z_k \subseteq \bigcup \{Z_c : c \in G - G'_k\}$; which implies that $\text{cl}(Z_{k\langle G'_k \rangle} - Z_k) \subseteq \bigcup \{Z_c : c \in G - G'_k\}$. Hence, $Z_k \cap \text{cl}(Z_{k\langle G'_k \rangle} - Z_k) \subseteq Z_k \cap \bigcup \{Z_c : c \in G - G'_k\} = \bigcup \{Z_k \cap Z_c : c \in G - G'_k\}$. Of course, $\#\bigcup \{Z_k \cap Z_c : c \in G - G'_k\} \leq \sum \{\#(Z_k \cap Z_c) : c \in G - G'_k\}$, and is thus $< N_1$ according to the definition of G'_k . Hence, $\#[Z_k \cap \text{cl}(Z_{k\langle G'_k \rangle} - Z_k)] < N_1$.

Suppose now there is some $i \in Z_{k\langle G'_k \rangle} - Z_k$, and take some $i' \in Z_k^\circ$ (which we have assumed not to be empty: $k \in G$). For some $\varepsilon > 0$, $D_{i'}(\varepsilon)$ is contained in Z_k° , and thus in $Z_{k\langle G'_k \rangle}$, which is convex; so that, for each $i'' \in D_{i'}(\varepsilon)$, the straight segment $[ii'']$ is contained in $Z_{k\langle G'_k \rangle}$. Hence, every point of $[ii'']$ either belongs to $Z_{k\langle G'_k \rangle} - Z_k$, or to Z_k (more precisely, to $Z_{k\langle G'_k \rangle} \cap Z_k$). As $i \in Z_{k\langle G'_k \rangle} - Z_k$ and $i'' \in Z_k$, some point of $[ii'']$ belongs to $Z_k \cap \text{cl}(Z_{k\langle G'_k \rangle} - Z_k)$. The number of the half-lines $[ii'')$ being clearly equal to N_1 , we have at least N_1 intersection points generated by segments like $[ii'']$; i.e., $\#[Z_k \cap \text{cl}(Z_{k\langle G'_k \rangle} - Z_k)] \geq N_1$ (in fact, it is equal to N_1 since $\#N^2 = N_1$), in contradiction with what we have said before. Hence, $Z_{k\langle G'_k \rangle} - Z_k = \emptyset$, so that $Z_k = Z_{k\langle G'_k \rangle}$.

Item (b) is now obvious; let us prove it concerning Z_k for instance. Under its hypotheses, if $k \leq c \forall c \in G'_k$, $Z_{k\langle G'_k \rangle}$ is convex and thus equal to Z_k because of item (a); moreover, this implies that Z_k° is convex because of Lemma 8.3. The proof of the sufficient condition ($\Leftarrow$) of item (c) is similar. Item (d) is a direct consequence of Prop. 2.7c or 5.6.c.iv.

Let us now prove the necessary condition ($\Rightarrow$) of item (c). First, concerning convexity. Suppose that Z_k is convex and consider some $c \in G'_k$. As a consequence of the definition of G'_k , $Z_k \cap Z_c$ has at least two points, i and i' . Those points thus belong to Z_k ; as Z_k is convex, $[ii'] \subseteq Z_k$. According to the hypotheses, $Z_k^\circ \neq \emptyset$; being open, that set contains some disk D with nonzero radius. Of course, $D - (ii') \neq \emptyset$; take some point p in that set. As Z_k

is convex, the whole triangle $[ii'p]$ is in Z_k , and thus also in Z_{kc} which contains Z_k , as shown by (6.7); and, as $p \notin (ii')$, $[ii'p]^\circ \neq \emptyset$.

Suppose now that $k \gg c$. The set $cl(Z_c^\circ)$ contains at least i and i' and is thus $\neq \emptyset$; hence, $Z_c^\circ \neq \emptyset$ [since $cl(\emptyset) = \emptyset$]; as $Z_c^\circ \subseteq Z_{ck}^\circ$, $Z_{ck}^\circ \neq \emptyset$. The strict convexifyingness of h thus implies that Z_{ck} is strictly convex, so that $]ii'[\subseteq Z_{ck}^\circ$. Thus, i'' being some point of $]ii'[$, $D_{i''}(\varepsilon) \subseteq Z_{ck}^\circ$ for some $\varepsilon > 0$. But $D_{i''}(\varepsilon) \cap [ii'p]^\circ \neq \emptyset$; hence, $Z_{ck}^\circ \cap Z_{kc}^\circ \neq \emptyset$, which is impossible (Prop. 5.7a): the proof is thus made *ab absurdo* that $k \leq c$.

The proof of the necessary condition ($\Rightarrow$) of item (c), concerning strict convexity, is easier. Now Z_k is strictly convex, so that $]ii'[\subseteq Z_k^\circ \subseteq Z_{kc}^\circ$; if $k \gg c$, Z_{ck}° is also strictly convex, and $]ii'[\subseteq Z_{ck}^\circ$. Hence, $Z_{ck}^\circ \cap Z_{kc}^\circ \neq \emptyset$ since it contains $]ii'[$; which is impossible. Thus, $k \leq c$. **Q.E.D.**

8.2. The three-centre problem

Concerning any three centres j , k , and q , contained in C , and any location i , we first have the obvious two implications that $(j \leq_i k \text{ and } k \leq_i q) \Rightarrow j \leq_i q$, and $j \leq_i q \Rightarrow (j \leq_i k \text{ or } k \leq_i q)$. Otherwise said,

$$Z_{qk} \cap Z_{kj} \subseteq Z_{qj} \subseteq Z_{qk} \cup Z_{kj} . \quad (8.1)$$

When $C = \{j, k, q\}$ and $q \leq k \leq j$, we have $Z_q = Z_{qj} \cap Z_{qk}$ and $Z_k = Z_{kj} \cap Z_{kq} = Z_{kj} - Z_{qk}^\circ$ (Prop. 5.7e). Fig. 8.1 shows the ensuing disposition of the market areas Z_k and Z_q when $Z_{qk} \cap Z_{kj} \neq \emptyset$.

Another clarifying property emerges from these equalities:

$$\begin{aligned} Z_k &= Z_{kj} \cup Z_{kq} \\ Z_j &= Z_{jk} \cup Z_{jq} \\ Z_{kj} &= Z_{qkj} \cup Z_{kqj} \cup Z_{kjq} \\ Z_{qj} &= Z_{kqj} \cup Z_{qjk} \cup Z_{qkj} , \end{aligned}$$

the latter two being instances of (6.11). They imply indeed that

$$Z_k \cup Z_q = Z_{kj} \cup Z_{qj} ;$$

which equality is visible on Fig. 8.1. In particular, when h is convexifying, Z_{kj} and Z_{qj} are convex; which allows us better to understand some properties of that case: that Z_k is convex when $Z_k \cap Z_q = \emptyset$ (because we have then $Z_k = Z_{kj}$), and that Z_j is arcwise connected when Z_k or Z_q is bounded.

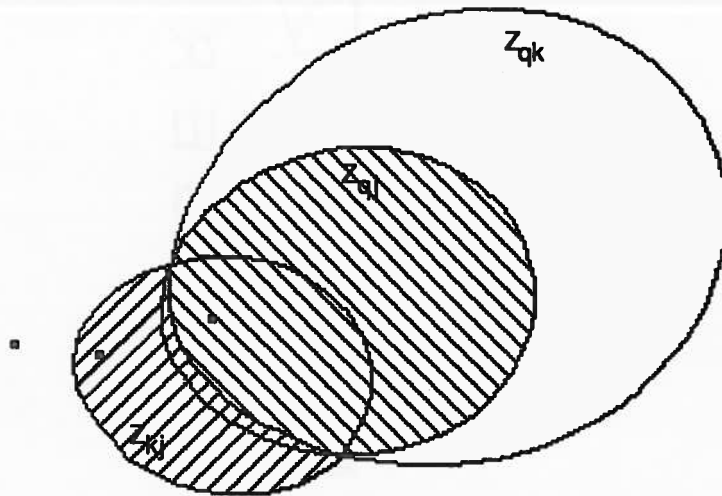


Fig. 8.1. — Market areas when $C = \{j, k, q\}$ with $j \gg k \gg q$; $h(\delta) = \delta^{0.5}$, $a = 1$, $a_k = 0.5$, $a_q = 0.25$.

Notice also that Z_j is of course arcwise connected when h is convex (Prop. 7.1c.ii), but that it is easy to build examples in which Z_j is not arcwise connected when the set of centres is $\{j, k, q\}$ as above, h not being convex, at least with functions like those used in Fig. 4.5 and 4.6.

We now come to a group of intermediate properties, the first of which is a kind of *duality principle*:

(8.2 – Duality principle) *If the points a and b belong to a same hypercircle having j and k as foci, j and k belong to a same hypercircle having a and b as foci.*

From the equalities $h_{aj} - h_{ak} = Q = h_{bj} - h_{bk}$, we deduce indeed that $h_{aj} - h_{bj} = h_{ak} - h_{bk}$, which yields the proof. **Q.E.D.**

Another elementary property will be useful:

(8.3) Consider three centres: j , k , and q .

(a) We have the equalities:

$$\begin{aligned} (Z_{jk} \cap Z_{kj}) \cap (Z_{jq} \cap Z_{qj}) &= (Z_{jk} \cap Z_{kj}) \cap (Z_{kq} \cap Z_{qk}) \\ &= (Z_{jq} \cap Z_{qj}) \cap (Z_{kq} \cap Z_{qk}); \end{aligned}$$

(b) if $C = \{j, k, q\}$, the three intersections of hypercircles listed sub (a) are equal to $Z_j \cap Z_k \cap Z_q$.

Those three intersections of hypercircles are indeed defined by the same inequalities: $h_j + p_j = h_k + p_k = h_q + p_q$; item (a) is thus obvious. As to item (b): according to (6.7), we have $Z_j = Z_{jk} \cap Z_{jq}$, $Z_k = Z_{kj} \cap Z_{kq}$, and $Z_q = Z_{qj} \cap Z_{qk}$; replacing Z_j , Z_k , and Z_q by those expressions in $Z_j \cap Z_k \cap Z_q$ shows that the latter set is equal to the common intersection of the three hypercircles. **Q.E.D.**

Let us now come back to the definition of strict convexifyingness, expressed in the situation studied in Part I: two centres, j and k . The t.c.f. h is strictly convexifying iff., for any $Q > 0$, if $Z_k^\circ \neq \emptyset$, Z_k is strictly convex. Since $Z_k \subseteq (jk)$ iff. $Z_k^\circ = \emptyset$, and since the indifference zone is the mediatrix of $\{j, k\}$ iff. $Q = 0$, the definition may be expressed in this alternative way: h is strictly convexifying iff. one of the two market areas is strictly convex or $Z_j \cap Z_k$ is equal to the mediatrix of $\{j, k\}$ or contained in (jk) .

The duality principle leads us to identify a new characteristic of maps when there are more than two centres. That property is also found, in a more general way, when the t.c.f. is *diskifying*; we define that h is such (w.r.t. our transportation space) iff., when two centres only are concerned, the market area of a less attractive centre is a disk. The functions $\ln$ and $\ln[(.)^2 + K]$, with $K > 0$, belong to that category. Of course, a diskifying t.c.f. is also strictly convexifying. But we first have to write down three lemmata.

LEMMA 8.4. *The intersection of a finite set of strictly convex sets is strictly convex.*

The lemma is obvious and its proof is omitted.

LEMMA 8.5. *Any subset E of the plane which is strictly convex and $E \neq \emptyset, \mathbb{R}^2$ contains no straight line.*

As $E \neq \emptyset$ and $E \neq \mathbb{R}^2$, both E and E^c are nonempty and $b(E)$ is nonempty. Of course, $b(E)$ contains at least two points, say a and b . As E is strictly convexifying, we have $]ab[\subseteq E^\circ$ and $(ab) - [ab] \subseteq E^c$. Suppose that a straight line Δ is contained in E . That line is thus necessarily $\neq (ab)$. If we draw the straight lines Δ_a and Δ_b // to Δ and respectively passing through a and b , it is clear that, E being convex, the strip S delimited by Δ_a and Δ_b is contained in E . Hence, $\Delta \neq (ab)$. But a consequence is that no point of S^c belongs to E . Otherwise, we might enlarge S to a strip S' delimited by Δ_a or Δ_b and the parallel to Δ passing through such a point; this would be impossible since a or b would belong to S'° and thus to E° . Hence, $E = S$; but S is not strictly convex. We thus arrive at a contradiction and there is no straight line contained in E . Q.E.D.

LEMMA 8.6. *If a subset E of the plane is closed and strictly convex, if a symmetry axis of E is chosen as the x -axis (the y -axis being perpendicular to the x -axis), and if the function f is defined as follows:*

$$\forall x \in \mathbb{R}: f(x) = \max \{y: (x, y) \in E\},$$

then f is continuous over $\text{dom}(f)$, and

$$b(E) \cap (\mathbb{R} \times \mathbb{R}_+) = \{(x, f(x)): x \in \text{dom}(f)\}$$

The last part of the statement, concerning $b(E)$, is obvious. Indeed, with those hypotheses, if $(x, y) \in E$, we have $(x, -y) \in E$, so that $b(E) \cap](x, y), (x, -y)[= \emptyset$. As to the other, suppose that f is not continuous in $x \in \text{dom } f$. There must be some $\varepsilon > 0$ for which, $\forall \eta > 0, \exists x' \in [x-\eta, x+\eta]$ with

$|y' - y| > \varepsilon$; we write y' for $f(x')$, as well as i for (x, y) , and i' for (x', y') . Two cases may be distinguished. Primo, if $y' > y + \varepsilon$, then the segment $[ii']$ and, because of symmetry, the whole trapeze $[i'i(x, -y)(x', y')]$ are contained in E ; hence, if we decrease η , the new point i' is necessarily also above $[ii']$. Secundo, if $y' < y - \varepsilon$, then, as again $[i'i(x, -y)(x', y')] \subseteq E$, we see that if we take a new η small enough, we are rejected to the first case. We so get to the conclusion that the segment $[(x, y)(x, y+\varepsilon)]$ belongs to $\text{cl}(E)$; i.e., to E itself since E is closed. Hence, $y \neq f(x)$: a contradiction. There is thus no discontinuity of f in any $x \in \text{dom}(f)$. **Q.E.D.**

Here is now the new property.

(8.4) (a) *When h is diskifying, the market areas of three centres have at most two points in common;*

(b) *when h is strictly convexifying, the market areas of three collinear centres, if their interiors are not empty, have at most one point in common on any of the two closed half-planes delimited by their straight line.*

Denote the three centres as j, k , and q , with $q \leq k \leq j$. Because of (6.7), we have $Z_j \cap Z_k \cap Z_q \subseteq Z_{j\langle j, k, q \rangle} \cap Z_{k\langle j, k, q \rangle} \cap Z_{q\langle j, k, q \rangle}$; and Prop. 8.3 tells us that the latter intersection is equal for instance to $(Z_{kj} \cap Z_{jk}) \cap (Z_{qj} \cap Z_{jq})$.

Let us begin with item (b). Assume the existence of two distinct points, a and b , in $Z_j \cap Z_k \cap Z_q$, and thus in the intersection of $Z_{kj} \cap Z_{jk}$ and $Z_{qj} \cap Z_{jq}$ for instance (Prop. 8.3), on a same closed half-plane limited by the straight line (jkq) . Suppose first that one point at least, point a for instance, does not belong to (jkq) . As a consequence of the duality principle (Prop. 8.2), j and k belong to some hypercircle with foci a and b , and so do q and j ; those two hypercircles are of course equal since they have j in common (that unique hypercircle is described by the equality $h_a - h_b = h_{ak} - h_{bk}$). But j, k , and q are collinear; as h is strictly convexifying, the straight line (jkq) is thus either the straight line (ab) , or the mediatrix of $\{a, b\}$. As such, it should pass either through a and b [discarded: a is supposed not to belong to (jkq) , precisely] or through m , the middle of $\{a, b\}$ [also

discarded: a and b being on the same side of $(j k q)$ and a lying out of it, so must be m]. We thus see, *ab absurdo*, that $a = b$.

Suppose now that both a and b belong to $(j k q)$. According to what we have just said, there is no other intersection point above the x -axis. According to Prop. 5.7d, $Z_{kj} \cap Z_{jk}$ and $Z_{qj} \cap Z_{jq}$ are the boundaries of Z_{kj} and Z_{qj} , respectively, since $q \leq k \leq j$. Those two areas have nonempty interiors, since they respectively contain Z_k and Z_j , which have been assumed to have nonempty interiors themselves. The t.c.f. being strictly convexifying, Z_{kj} and Z_{qj} are thus strictly convex; so that, for every x , according to Lemma 8.6, there is at most one point (x, y) for which $y \geq 0$ in $Z_{kj} \cap Z_{jk}$; and similarly for $Z_{qj} \cap Z_{jq}$. Those points, if any, are $(x, f_{kj}(x))$ and $(x, f_{qj}(x))$, respectively; the functions $f_{cc'}$ are so defined:

$$f_{cc'}(x) = \max \{y; (x, y) \in Z_{cc'}\} . \quad (8.2)$$

The set on the right side may be empty, so that $f_{cc'}(x) = \max \emptyset = -\infty$; otherwise, $f_{cc'}(x) \geq 0$. We know from Lemma 8.6 that f_{kj} and f_{qj} are continuous on their definition sets; both of those sets are equal to the interval $[x_a, x_b]$ since Z_{kj} and Z_{qj} are strictly convex.

As we have just said, the graphs of those functions do not cross when $y > 0$; hence, because they are continuous, we have either, $\forall x \in]a, b[, f_{kj}(x) < f_{qj}(x)$, or $f_{qj}(x) < f_{kj}(x)$. Assume for instance that the latter property holds; i.e., $Z_{qj} \subset Z_{kj}$. Hence we may find a point, say i , in the intermediate region $Z_{kj}^\circ - Z_{qj}$. Because h is continuous, it is possible to increase slightly a_q in such a way that point i remains in $Z_{kj}^\circ - Z_{qj}$ but $a \in Z_{qj}^\circ$ (indeed, initially, $q \neq k$: otherwise we should have $Z_{kj} = Z_{qj}$). We have thus now $f_{kj}(x_a) = 0 < f_{qj}(x_a)$ but $f_{qj}(x_i) < f_{kj}(x_i)$: the graphs of the two functions must now cross for some $x \in]x_a, x_i[$, a fact which we know to be impossible from the first part of the present proof of item (b).

The proof of item (a) is thus immediate: function h being diskifying, $Z_{kj} \cap Z_{jk}$ and $Z_{qj} \cap Z_{jq}$ are either circles with nonzero radius, or straight lines (when the corresponding two centres are equally attractive), or points. Hence, they have at most two points in common, unless they are equal. But the latter case is to be discarded: 1° if they are equal to a straight line, then $j = k = q$, and $Z_{kj} \cap Z_{jk} = Z_{qj} \cap Z_{jq}$ iff. $k = q$; 2° if the two hypercircles are

the same circle, then k and q are on the straight line (jo) , in which o is the centre of the circle, and we are in the case treated in item (b); and 3° if $Z_{kj} \cap Z_{jk}$ and $Z_{qj} \cap Z_{jq}$ are reduced to one point, they have of course not three points in common. We have thus $\#[(Z_{kj} \cap Z_{jk}) \cap (Z_{qj} \cap Z_{jq})] \leq 2$; from which $\#(Z_j \cap Z_k \cap Z_q) \leq 2$, since we have just seen that $Z_j \cap Z_k \cap Z_q \subseteq (Z_{kj} \cap Z_{jk}) \cap (Z_{qj} \cap Z_{jq})$. **Q.E.D.**

To be complete, let us notice here that, in the light of the subsequent Prop. 8.5, the proof of Prop. 8.4b will appear as a particular case of a general property that can be directly applied to interiors of market areas. (As a matter of fact, strictly speaking, we should term that proposition a conjecture, since we accept one point of the proof without theoretical basis, just because it is obvious, as will be seen). Before that, however, we need to establish a lemma. We say that the path P links the points a and b through E iff. a and b are extremities of P and $P - \{a, b\} \subseteq E$.

LEMMA 8.7. *If the set E is arcwise connected, any points a and b of $b(E)$ may be linked by some path P through E .*

The points a and b adhering to E , we can find for any $\varepsilon_0 > 0$ two points in E , a_0 and b_0 , for which $\delta_{aa_0} \leq \varepsilon_0$ and $\delta_{bb_0} \leq \varepsilon_0$. We similarly define, for any $n \in \mathbb{N}$, a_n and b_n as points in E for which $\delta_{aa_n} \leq \varepsilon_n$ and $\delta_{bb_n} \leq \varepsilon_n$; with for instance $\varepsilon_n = \varepsilon_0/2^n$. The set E being arcwise connected, there is a path P_0 linking a_0 to b_0 and contained in E . Then, we may link a_0 to a_1 and b_0 to b_1 by paths that are contained in E ; the union of those paths and of P_0 forms a new path, P_1 . We so construct a series of paths $(P_n: n \in \mathbb{N})$, with $P_n \subseteq P_{n+1}$ and the extremities of which, a_n and b_n , get closer and closer to a and b , respectively. It is thus clear that that series tends toward a path P_∞ linking a and b through E . **Q.E.D.**

Here is now the general property.

(8.5) *If A , B , and C are three open arcwise connected sets of the plane, with $A \cap B = B \cap C = C \cap A = \emptyset$, then $\#[cl(A) \cap cl(B) \cap cl(C)] \leq 2$.*

First, notice that $A \cap b(B) = \emptyset$. Otherwise, there would be some point $i \in b(B)$ and some $\varepsilon > 0$ for which $D_i(\varepsilon) \subseteq A$; but point i would at the same time adhere to B , so that $D_i(\varepsilon)$ would intersect B ; hence, $A \cap B \neq \emptyset$: a contradiction with the hypotheses. Thus none of the sets A , B , and C has a nonempty intersection with the boundary of any other of those sets. Hence, $\text{cl}(A) \cap \text{cl}(B) \cap \text{cl}(C) = b(A) \cap b(B) \cap b(C)$; see indeed (5.5).

Suppose there are three points, a , b , and c , in that intersection; see Fig. 8.2, in which we have used, for clarity's sake, paths containing no loops. As A , B , and C are arcwise connected, we can draw three paths P_A , P_B , and P_C linking a to b through A , B , and C , respectively, as indicated by Lemma 8.7. Since $A \cap B = B \cap C = C \cap A = \emptyset$, we have $P_A \cap P_B = P_B \cap P_C = P_C \cap P_A = \{a, b\}$. The set $\{i: \exists \Gamma: \Gamma \text{ is a path linking } a \text{ and } b, i \in \Gamma, \text{ and } \Gamma \cap P_A = \Gamma \cap P_B = \Gamma \cap P_C = \emptyset\}$ has three arcwise connected components that we denote as $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$; P_A is isolated from $\mathcal{A}$ by $P_B \cup P_C$ [i.e., any path joining a point of P_A to a point of $\mathcal{A}$ meets $P_B \cup P_C$], P_B is isolated from $\mathcal{B}$ by $P_C \cup P_A$, P_C is isolated from $\mathcal{C}$ by $P_A \cup P_B$ ².

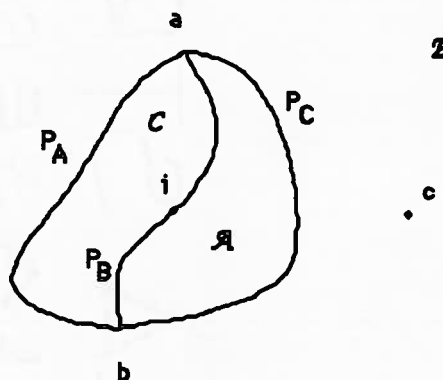


Fig. 8.2. — Proof of Prop. 8.5.

The point c is distinct from a and b , does not belong to P_A , P_B or P_C (since it does not belong to A , B or C), and may be linked to a and b (through A , B , and C); it is thus in $\mathcal{A}$, $\mathcal{B}$, or $\mathcal{C}$. Suppose for instance that it belongs to $\mathcal{B}$. Take now some point $i \in P_B - \{a, b\}$. According to Lemma 8.7, there is some path P linking c to i through B ; but i and c are isolated from one

² I admit the properties expressed in that sentence without theoretical basis.

another by $P_A \cup P_C$. This is impossible since $P - \{c\}$, contained in B , cannot intersect P_A nor P_C . We thus arrive at a contradiction and the proof is made *ab absurdo*. **Q.E.D.**

8.3. Arcwise-connectedness

We begin this section with a general statement which is similar to the preceding Prop. 8.1a and extends Prop. 7.3:

(8.6) If $k \in G$ and if $Z_k \langle G_k \rangle$ [resp. $Z_k \langle G_k^{(0)} \rangle$] is arcwise connected, then
 $Z_k = Z_k \langle G_k \rangle$ [resp. $Z_k^o = Z_k^o \langle G_k^{(0)} \rangle$].

The proof is completely similar to that of Prop. 7.3. We have just to replace 'starshaped w.r.t. k ' by 'arcwise connected' and the segment $[ik]$ by some path P linking i to some $i' \in Z_k$. **Q.E.D.**

When a market area is starshaped or convex, it is necessarily arcwise connected; but the reverse is not true. A market area may be arcwise connected without being starshaped nor convex. But it seems rather difficult to obtain a market area that is not arcwise connected when h is concave or convex; see Thill's work, for instance, where that case never occurs. The possibility exists, nevertheless, as shown in Fig. 8.3.

In some cases, it is possible to predict the arcwise-connectedness of Z_k when there are more than two centres, as we have seen in the beginning of Section 8.2 concerning three centres. Another such property, close to Prop. 8.4b, is the following:

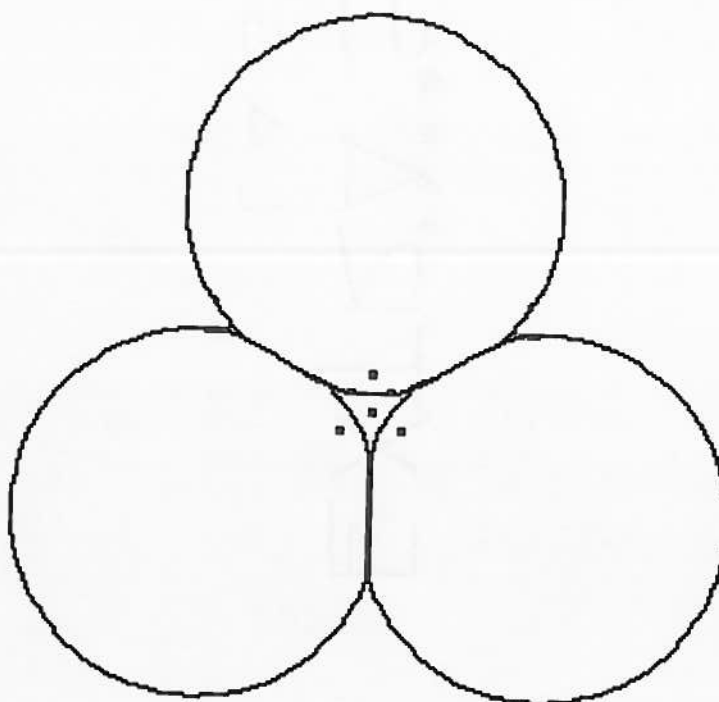


Fig. 8.3. — Market areas when there are 4 centres. The most attractive centre, with attractiveness 1.1, is in the triangle formed by the three others, with attractivenesses equal to 1.

- (8.7) If h is convexifying, if all the centres are collinear, if $j \in G$, and if j is most attractive among G_j [resp. $G_j^{(o)}$], Z_j^o and Z_j are arcwise connected [resp. $Z_j^{(o)}$ is arcwise connected].

(We give the proof when $G_j = G_j^{(o)} = G$; the rest is easily deduced from Prop. 8.6). Let us assimilate that straight line with the x -axis, and construct a y -axis perpendicular to it. The functions f_{cj} have been defined in (8.2). If, for some c and some x , $f_{cj}(x) = \infty$, then, as Z_{cj} is convex (because h is convexifying) and symmetric w.r.t. the x -axis, the whole vertical line $\{x\} \times \mathbb{R}$ is contained in Z_{cj} . As Z_{cj} is convex and does not contain j , this is possible only if Z_{cj} is a half-plane having a vertical boundary. Denote by G' the centres $c \in G$ for which Z_{cj} has the latter property. Define now

$$f(x) = \max \{f_{cj}(x) : c \in G - G'\} ;$$

this function thus never takes the value ∞ .

As a consequence of (6.7),

$$Z_j^\circ = \bigcap \{Z_{jc}^\circ : c \in G - \{j\}\} = Z_j^\circ \langle G' \cup \{j\} \rangle \cap Z_j^\circ \langle G - G' \rangle .$$

The set $Z_j^\circ \langle G' \cup \{j\} \rangle$, equal to $(\bigcup \{Z_{cj} : c \in G'\})^-$ because of Prop. 5.8, is either an open half-plane parallel to the y-axis, or a vertical open strip parallel to the y-axis; or the plane when $G' = \emptyset$. As to $Z_j^\circ \langle G - G' \rangle$, similarly equal to $(\bigcup \{Z_{cj} : c \in (G - \{j\}) - G'\})^-$, it is clear that it is arcwise connected. Take indeed two points i and i' of that set above the x-axis; thus $y_i > f(x_i)$ and $y_{i'} > f(x_{i'})$. As f does not take the value ∞ between x_i and $x_{i'}$, it is of course possible to join i and i' through a path remaining in $Z_j^\circ \langle G' \cup \{j\} \rangle$; for instance, if $x_i < x_{i'}$, and if we set $\mu_{ii'} = \max\{0, \max \{f(x) : x_i \leq x \leq x_{i'}\}\}$, the path may be

$$[i(x_i, \mu_{ii'})] \cup [(x_i, \mu_{ii'})(x_{i'}, \mu_{ii'})] \cup [(x_{i'}, \mu_{ii'}) i'] .$$

That particular path is of course also contained in $Z_j^\circ \langle G' \cup \{j\} \rangle$ if i and i' additionally belong to $Z_j^\circ \langle G' \cup \{j\} \rangle$. Consequently, Z_j° is arcwise connected above the x-axis; and as j , which belong to the x-axis, also belongs to Z_j° (Prop. 7.6a.ii), Z_j° is arcwise connected. The proof of the arcwise-connectedness of Z_j is similar and left to the reader. **Q.E.D.**

Even when the other market areas are convex, it is quite possible, however, that the market area of the most attractive centre is not arcwise connected. We have seen this in the beginning of Section 8.2 with three centres, Z_q and Z_k being unbounded, concerning Fig. 4.6; and in Fig. 8.3.

Concerning again collinear centres, Prop. 8.4b leads us to this statement:

(8.8) Suppose that all the centres are collinear and that the t.c.f. is strictly convexifying.

(a) Every market area either is arcwise connected, or has two arcwise connected components symmetric of one another w.r.t. the line of centres; the same is true for the interior of the market area;

(b) if the t.c.f. h is concave, all the market areas are arcwise connected; if $h(0) = -\infty$, all the market areas and their interiors are arcwise connected.

Let us consider the centres' straight line as the x-axis; the y-axis is perpendicular to it. The market area of any centre, call it k , is equal to $\bigcap \{Z_{kc} : c \in C\}$. We may thus write

$$Z_k = Z_{k\leftarrow} \cap Z_{k\rightarrow} \cap Z_{k\equiv}, \quad (8.3)$$

with

$$Z_{k\leftarrow} = \bigcap \{Z_{kc} : k \leftarrow c\}$$

$$Z_{k\rightarrow} = \bigcap \{Z_{kc} : c \leftarrow k\}$$

$$Z_{k\equiv} = \bigcap \{Z_{kc} : k \equiv c\}.$$

The last of those sets is easy to compute. If there are centres $\equiv k$ on the right and on the left of k ,

$$Z_{k\equiv} = \left[\frac{x_j + x_k}{2}, \frac{x_k + x_q}{2} \right] \times \mathfrak{R},$$

where j is, among those centres, the centre closest to k on the left of k , and q is the centre closest to k on its right. Supposing that $\#C > 1$, if there is no centre $\equiv k$ on the left of k , $Z_{k\equiv} =]-\infty, (x_j + x_k)/2] \times \mathfrak{R}$; if there is no centre $\equiv k$ on its right, $Z_{k\equiv} = [(x_j + x_k)/2, -\infty[\times \mathfrak{R}$. Anyway, $Z_{k\equiv}$ is thus an interval, that we denote as I in the following.

As to the two other sets in (8.3), notice first, as in Prop. 8.4b, that any 'vertical' line (i.e., parallel to the y-axis) cuts $Z_{ck} \cap Z_{kc}$ in the half plane above the x-axis ($y \geq 0$) in at most one point when $k \leftarrow c$ or $k \rightarrow c$. Given

the abscissa x of that vertical line, the point, if any, is (x, y) , with $y = f_{ck}(x)$ if $c \ll k$, or $y = f_{kc}(x)$ if $k \ll c$; see (8.2). If there is no intersection point, we so find $y = -\infty$.

It so appears that we may write:

$$Z_{k\ll} = \{(x, y) : x \in \mathfrak{R} \text{ and } |y| \leq \min \{f_{kc}(x) : k \ll c\}\}$$

$$Z_{k\gg} = \{(x, y) : x \in \mathfrak{R} \text{ and } |y| \geq \max \{f_{ck}(x) : k \gg c\}\},$$

and thus also (supposing for instance the presence of centres on both sides of k)

$$Z_k \cap (\mathfrak{R} \times \mathfrak{R}_+) =$$

$$\{(x, y) : x \in I, y \geq 0 \text{ and } f_{k-}(x) \leq y \leq f_{k+}(x)\},$$

in which we have set

$$f_{k+}(x) = \min \{f_{kc}(x) : k \ll c\}$$

$$f_{k-}(x) = \max \{f_{ck}(x) : k \gg c\}.$$

Let us now have a closer look at those two functions. As said above, when $k \ll c$, $f_{kc}(x) < \infty$; thus, $f_{k+}(x) < \infty$. On the other hand, it is clear that f_{k+} and f_{k-} are continuous on their definition set, because this is also true for the functions f_{kc} (if $k \ll c$) and f_{ck} (if $c \ll k$) as also said above. It is also clear that $\text{dom}(f_{kc})$ (if $k \ll c$), or $\text{dom}(f_{ck})$ (if $c \ll k$), is an interval, since Z_{kc} or Z_{ck} , respectively, is convex. Hence, $\text{dom}(f_{k+})$ is an interval, being an intersection of intervals. Outside its domain, f_{k+} takes the value $-\infty$.

Let us now show that f_+ and f_- cannot have more than two intersection points. Assume the existence of such a point, i . There are some centres c and c' such that $k \ll c$ and $f_+(x_i) = f_{kc}(x_i)$, and $k \gg c'$ and $f_-(x_i) = f_{c'k}(x_i)$. According to Prop. 8.4b, the graphs of the functions f_{kc} and $f_{c'k}$ do not cross for any other value of x . Since 1° the graphs may only intersect in the intersection of the domains of both functions, and 2° those functions are continuous on their respective domains and thus in their intersection, which is moreover an interval, we have either, $\forall x \in \text{dom}(f_{kc}), f_{kc}(x) < f_{c'k}(x)$ for $x < x_i$, and $f_{kc}(x) > f_{c'k}(x)$ for $x > x_i$; or vice versa.

Hence, we see that, according to the definition of f_+ and f_- , we have, either on $\text{dom}(f_+) \cap]-\infty, x_i[$ or on $\text{dom}(f_+) \cap]x_i, \infty[$,

$$f_+(x) \leq f_{kc}(x) < f_{c'k}(x) \leq f_-(x),$$

and thus $f_+(x) < f_-(x)$. I.e., either the graphs of the functions f_+ and f_- do not cross on the left of x_i , or they do not on its right. We suppose from now on that we are in the former of the two possibilities; the other may be treated similarly.

A first case is when, still on $\text{dom}(f_+)$, there is no intersection point on the right of x_i either. For every $x \in \text{dom}(f_+) \cap [x_i, \infty[$, the points of Z_k intercepted by the vertical half-line $\{x\} \times \mathbb{R}_+$ form the segment $\{x\} \times [\max\{0, f_-(x)\}, f_+(x)]$. That segment of course intercepts the graph of the function

$f_{\pm}$, defined as $(\max\{0, f_-\} + f_+)/2$. The function $f_{\pm}$ is in this case $> \max\{0, f_-\}$ and $< f_+$ on $\text{dom}^\circ(f_+) \cap]x_i, \infty[$; at the right extremity (if any) of the interval $\text{dom}(f_+)$, the three functions take the same value, 0. As the graph of $f_{\pm}$ is arcwise-connected because f_+ and $\max\{0, f_-\}$ are continuous on $\text{dom}(f_+)$, which is also $\text{dom}(f_{\pm})$, we see that $Z_k \cap (\mathbb{R} \times \mathbb{R}_+)$ and $Z_k^\circ \cap (\mathbb{R} \times \mathbb{R}_+)$ are arcwise connected.

But the fact that the graphs of f_+ and f_- do not intersect at all on $\text{dom}(f_+) - \{x_i\}$ and that f_+ takes 0 as value in the extremities of $\text{dom}(f_+)$ implies that those extremities, if any, are outside $\text{dom}(f_-)$. The set $\text{dom}(f_-)$ is a union of closed intervals, $\cup\{\text{dom}(f_{c'k}) : c' \ll k\}$; it is thus closed itself, and each extremity of $\text{dom}(f_+)$ thus belongs to an open set on which f_- takes the value $-\infty$. Which means that for every x in some interval around each such an extremity, the point $(x, 0)$ is in Z_k . As $\text{dom}(f_+)$ has at least one extremity (otherwise, $Z_{k\ll}$ would contain the x -axis and could not be strictly convex; see Lemmata 8.4 and 8.5), Z_k and Z_k° are thus arcwise connected.

The second case is when there is a second intersection point, i' , at the right of i . Of course, $f_- < f_+$ on $]x_i, x_{i'}[$. Here again, we must have, for some centres p and p' and for all $x \in \text{dom}(f_+)$,

$$\forall x > x_i : f_+(x) \leq f_{kp}(x) < f_{kp'}(x) \leq f_-(x).$$

There is thus no other intersection point of the graphs of f_- and f_+ . Here again, also, the graph of the function $(\max\{0, f_- \} + f_+)/2$ is arcwise connected and so are $Z_k \cap (\mathbb{R} \times \mathbb{R}_+)$ and $Z_k^\circ \cap (\mathbb{R} \times \mathbb{R}_+)$.

Item (a) is thus obvious. Item (b) is now also evident: the additional conditions ensure that $k \in Z_k$. As k is on the x -axis, it belongs to the intersection of Z_k with the two closed half-planes delimited by that axis; both being arcwise connected, every point of Z_k can be joined to k through some path; Z_k itself is thus arcwise connected. Moreover, if $h(0) = -\infty$, $k \in Z_k^\circ$ (h being continuous *lato sensu* in 0), which shows the rest. Q.E.D.

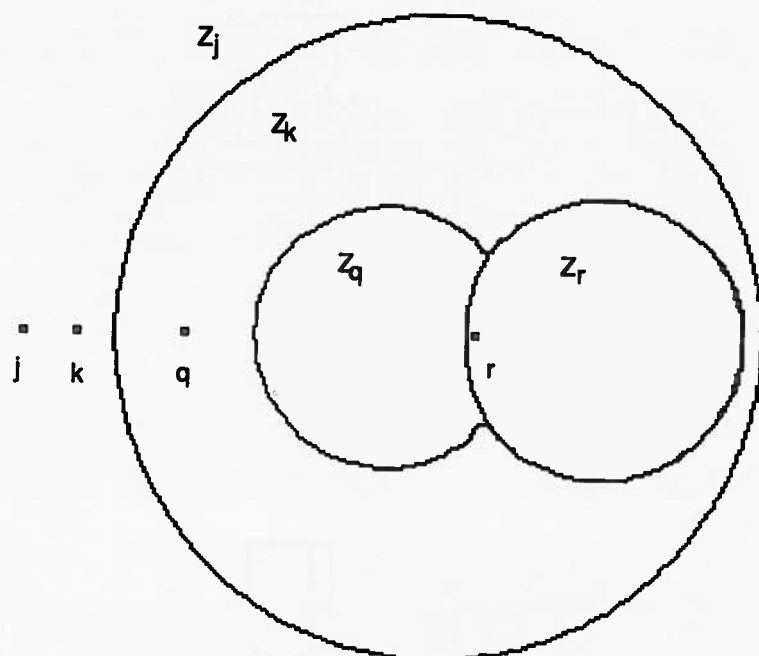


Fig. 8.4. — Illustration of Prop. 8.8: the centres j , k , q , and r are collinear;
 $a_j = 7$, $a_k = 6.1$, $a_q = 4$, $a_r = 1.45$; $x_j = -2$, $x_k = 0$, $x_q = 4$, $x_r = 15$; $h = \ln[(.)^2 + 64]$.

All market areas are arcwise connected.

We give in Fig. 8.4 an illustration of that proposition, concerning the function $\ln [(.)^2 + K^2]$, which is not concave, but strictly convexifying since the intermediate zone is a circle when two centres only are concerned (Prop. 2.7e).

8.4. Choice and preference areas

Concerning convexity and choice areas, we may just say this. If we take $m = \#C - 1$, (6.9) rewrites as

$$Ch_k(\#C - 1) = \bigcup \{Z_{kc} : c \in C - \{k\}\} ;$$

hence,

$$Ch_k^-(\#C - 1) = \bigcap \{Z_{kc}^- : c \in C - \{k\}\} .$$

Thus, if k is most attractive, $Z_{kc}^- = Z_{ck}^0$ (Prop. 5.7e), which is convex if h is convexifying. In that case, $Ch_k^-(\#C - 1)$ is convex.

As to preference areas, it is clear from (6.10) that $Z_{c_1 c_2 \dots c_n}$ is convex if h is convexifying and $c_m \leq c_{m+1}$, $\forall m \in [1, m-1] \cap \mathbb{N}$; it is strictly convex if h is strictly convexifying and $c_m < c_{m+1}$, $\forall m \in [1, m-1] \cap \mathbb{N}$.

If all the centres are on a same straight line and h is convexifying, the proof of Prop. 8.7 and (6.9) show that the areas $Ch_j(m)$, for all integer $m \in [1, m[$, are arcwise connected if j is most attractive. Prop. 8.8b and its proof imply that, when h is strictly convexifying, the areas $Ch_k(m)$, k being now any centre, are arcwise connected if h is concave or $h(0) = -\infty$.

9. Other theoretical results

In this chapter, we shall make use of a well-known property of convex (or, *mutatis mutandis*, concave) functions in order to explore dominance areas and total potential a step further. Given a real function f , a list of real numbers $(\xi_i: i \in [1, m] \cap \mathbb{N})$, and a list of weights $(\lambda_i: i \in [1, m] \cap \mathbb{N})$, with $\sum(\lambda_i: i \in [1, m] \cap \mathbb{N}) = 1$, the weighted arithmetic average of the couples $(\xi_i, f(\xi_i))$, with abscissa $\mu = \sum(\lambda_i \xi_i: i \in [1, m] \cap \mathbb{N})$ lies in the convex hull of their set. Hence, if f is a convex function, $f(\mu) \leq \sum(\lambda_i f(\xi_i): i \in [1, m] \cap \mathbb{N})$; the property is illustrated in Fig. 9.1.

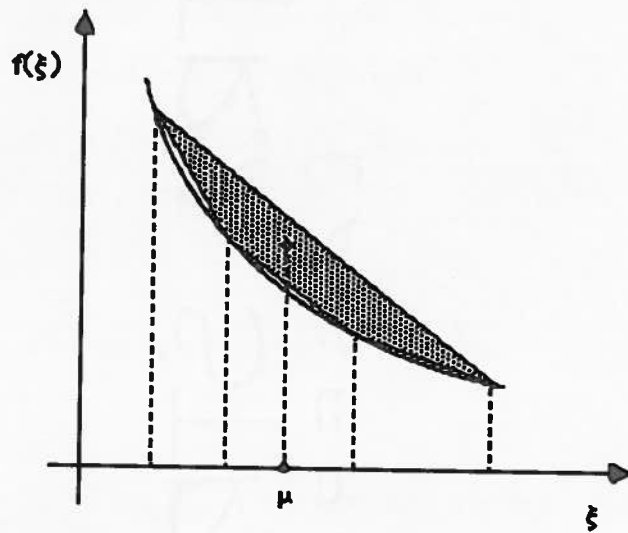


Fig. 9.1. — If a function is convex, the image it gives of a weighted arithmetic average of numbers is not larger than the corresponding weighted arithmetic average of images.

9.1. Approximations of dominance areas

How can we go further in the study of Y_k ? One way is to find an approximate area easier to study or to draw. Consider some point, p , of the plane. The quantity s_k , similarly to (7.5), is $\sum(a_c: c \in C - \{k\})$; and let

$$\lambda_{kc} = a_c / s_k ,$$

$$\omega_{kp} = \sum_{c \in C - \{k\}} \lambda_{kc} \delta_{pc} .$$

The following area may be thought of as an approximation of Y_k :

$$\mathcal{E}_{kp} = \{i: a_k F_k \geq s_k F(\delta_p + \omega_{kp})\} ,$$

as we shall see through the proof of the first item of the next proposition; whereas the second shows that $\mathcal{E}_{kp}$ is at least easier to draw than Y_k . Before reading the next proposition, remember these implications:

h is concave $\Rightarrow F$ is convex

F is concave $\Rightarrow h$ is convex

(9.1) (a) When F is convex, $Y_k \subseteq \mathcal{E}_{kp}$;

(b) When k is dominant but not strictly:

(i) if $\delta_{kp} > \omega_{kp}$, $\mathcal{E}_{kp} = \emptyset$;

(ii) if $\delta_{kp} \leq \omega_{kp}$, $\mathcal{E}_{kp}$ is a concave area containing p and limited by the curve $\delta_k - \delta_p = \omega_{kp}$, which is a branch of a hyperbola with foci k and p ;

(c) When h is concave:

(i) $\mathcal{E}_{kp} \neq \emptyset \Leftrightarrow k \in \mathcal{E}_{kp}$; $\mathcal{E}_{kp}^- \neq \emptyset \Leftrightarrow p \in \mathcal{E}_{kp}^-$;

(ii) if k is not strictly dominant, $\mathcal{E}_{kp}$ is starshaped w.r.t. k ;

(iii) if k is dominant, $\mathcal{E}_{kp}^-$ is starshaped w.r.t. p ;

(d) When h is convex:

(i) $\mathcal{E}_{kp} \neq \emptyset \Leftrightarrow \mathcal{E}_{kp}$ is unbounded; $\mathcal{E}_{kp}^- \neq \emptyset \Leftrightarrow \mathcal{E}_{kp}^-$ is unbounded;

(ii) if k is not strictly dominant, $\mathcal{E}_{kp}^-$ is starshaped w.r.t. p ;

(iii) if k is dominant, $\mathcal{E}_{kp}$ is starshaped w.r.t. k .

Item (b) is plain. The items (c) and (d) are obvious once we extend our transportation space to a new one, $\mathcal{R}^2 \cup \{p'\}$, where p' , a point out of $\mathcal{R}^2$, may be linked to $\mathcal{R}^2$ only through p (so that the starshapedness of a subset of $\mathcal{R}^2$ w.r.t. p' implies starshapedness w.r.t. p), with $\delta_{pp'} = \omega_{kp}$. This allows us to rewrite the definition of $\mathcal{E}_{kp}$ as follows:

$$\mathcal{G}_{kp} = \{i: a_k F_k \geq s_k F_p\} ,$$

to which we may apply Prop. 5.4, since that proposition is independent of the transportation space.

As to item (a), the following equality:

$$\sum_{c \in C - \{k\}} a_c F_c = s_k \sum_{c \in C - \{k\}} \lambda_{kc} F_c , \quad (9.1)$$

makes appear a linear combination of the F_c 's with positive coefficients λ_{kc} ; and

$$\sum_{c \in C - \{k\}} \lambda_{kc} = 1 .$$

If F is convex, the property described in Fig. 9.1 implies that

$$\sum_{c \in C - \{k\}} \lambda_{kc} F_c \geq F \left(\sum_{c \in C - \{k\}} \lambda_{kc} \delta_c \right) . \quad (9.2)$$

On the other hand, we have the triangle inequality:

$$\delta_c \leq \delta_p + \delta_{pc} ,$$

from which

$$\begin{aligned} & \sum_{c \in C - \{k\}} \lambda_{kc} \delta_c \\ & \leq \sum_{c \in C - \{k\}} \lambda_{kc} (\delta_p + \delta_{pc}) = \delta_p + \sum_{c \in C - \{k\}} \lambda_{kc} \delta_{pc} = \delta_p + \omega_{kp} . \end{aligned}$$

The deterrence function F being decreasing, we have thus

$$F \left(\sum_{c \in C - \{k\}} \lambda_{kc} \delta_c \right) \geq F (\delta_p + \omega_{kp}) ,$$

which, together with (9.1) and (9.2), shows that

$$\sum_{c \in C - \{k\}} a_c F_c \geq s_k F (\delta_p + \omega_{kp}) .$$

Hence, if a point i belongs to Y_k , $a_k F_k \geq s_k F(\delta_p + \omega_{kp})$, so that $i \in \mathcal{G}_{kp}$.
Q.E.D.

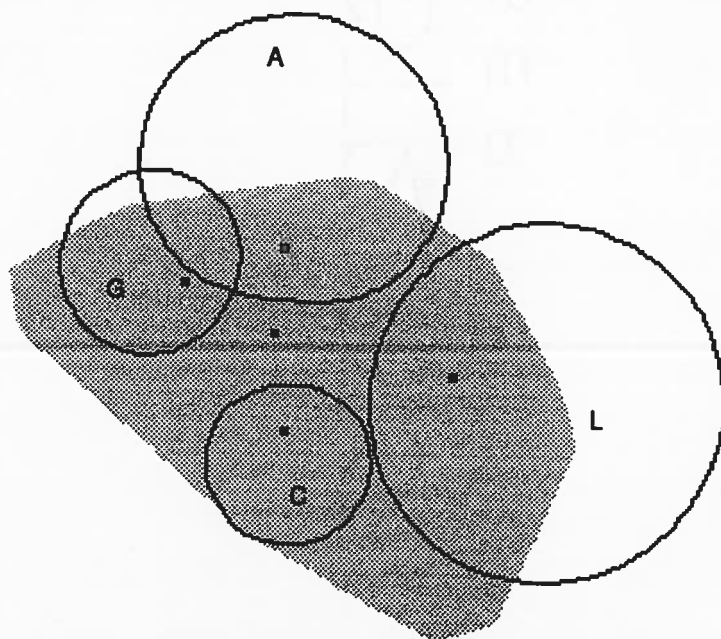


Fig. 9.2. — Map of the areas $\mathcal{C}_{k,w}$ associated with Belgian cities of the third level of urban hierarchy in Belgium; see Fig. 6.5.

The Weber point of Brussels is Antwerp;
the other cities have Brussels as Weber point.
No boundary appears for Brussels.

A question remains: how should we choose the point p ? The main distortion being probably introduced by replacing δ_c by $\delta_p + \delta_{pc}$, it seems reasonable to minimize the weighted sum of the terms δ_{pc} ; i.e., ω_{kp} . The point which achieves that goal is well known: it is the *Weber point*, k' , corresponding to the set $C - \{k\}$; each element of which, c , being weighted by a_c .

Another choice of p may be of interest: if we set $p = k$, the inequality defining $\mathcal{C}_{kp}$ becomes:

$$a_k F(\delta_k) \geq s_k F(\delta_k + \omega_{kk}) ,$$

i.e.,

$$h(\delta_k + \omega_{kk}) - h(\delta_k) \geq \ln(s_k / a_k) .$$

Hence appears the next property:

(9.2) (a) If k is dominant, $\mathcal{E}_{kk} = \mathcal{R}^2$;

(b) If h is concave, $\mathcal{E}_{kk}$ is a disk centred on k ;

(c) If h is convex, $\mathcal{E}_{kk}$ is the exterior of an open disk centred in k ;

(d) If h is linear, $\mathcal{E}_{kk}$ is either $\mathcal{R}^2$ or $\emptyset$.

In particular, when F is the Pareto function $\delta^{-\beta}$, the radius of $\mathcal{E}_{kk}$ is:

$$R_k = \omega_{kk} a_k^{1/\beta} / (s_k^{1/\beta} - a_k^{1/\beta}) .$$

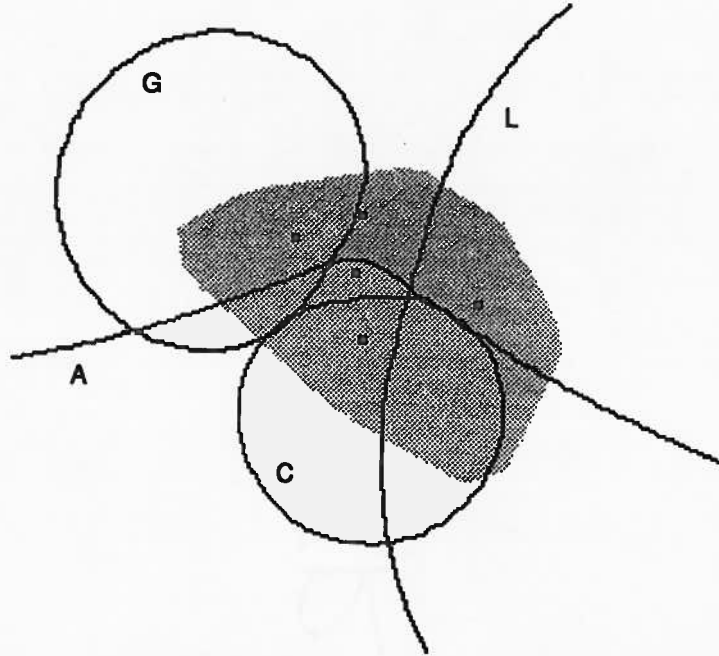


Fig. 9.3. — Approximations of arithmetic average areas by areas of type $\mathcal{E}_{kw}$; see Fig. 6.6.

No boundary appears for Brussels.

Let us now come to a second bound, that writes

$$\mathcal{E}_k = \{i: a_k F_k \geq s_k F [(\delta_k^{*2} + \rho_k^2)^{1/2}]\} ,$$

where k^* is the inertia centre of the set $C - \{k\}$ when the mass associated with each centre is its attractiveness; ρ_k is the corresponding inertia radius.

We also denote by g the inertia centre of the set C when each centre, c , receives a mass a_c . We have thus

$$\sum_{c \in C - \{k\}} \lambda_{kc} \delta_c^2 = \delta_k^2 + \rho_k^2. \quad (9.3)$$

Note these implications, in addition to those mentioned above:

$$\begin{aligned} F \circ \sqrt{\cdot} \text{ is convex} &\Rightarrow F \text{ is convex} \\ F \text{ is concave} &\Rightarrow F \circ \sqrt{\cdot} \text{ is concave} \end{aligned}$$

The following properties obtain:

- (9.3) (a) (i) If $F \circ \sqrt{\cdot}$ is convex, $Y_k \subseteq E_k$;
(ii) if $F \circ \sqrt{\cdot}$ is concave on the range of distances considered, $E_k \subseteq Y_k$;
(b) (i) If h is concave: $k \in E_k \Leftrightarrow E_k \neq \emptyset$;
(ii) if h is concave and k is not strictly dominant, or if h is convex and k is dominant, or if h is linear, E_k is starshaped w.r.t. k ;
(c) When h is convexifying [resp. strictly convexifying]:
(i) if k is not dominant, E_k is convex [resp. strictly convex];
(ii) if k is strictly dominant, $cl(E_k^-)$ is convex [resp. strictly convex];
(d) If $h = \ln$, or if $h = \ln[(\cdot)^2 + K]$ with $K > 0$, $b(E_k)$ is a circle or a point, or is empty;
(e) If k is dominant but not strictly: $b(E_k)$ is a straight line perpendicular to (kk') , closer to k' than to k , and the distance between it and the mediatrix of $[kk']$ is $\rho_k^2 / (2 \delta_{kk'})$;
(f) $g \in (kk')$.

The proof of (a) is immediate: see (9.3) above. Item (f) is obvious, since g is the inertia centre of $\{k, k'\}$, k receiving a mass a_k and k' a mass s_k . As to the other items, we have simply to extend the transportation space to $\mathbb{R}^3$ and consider the point k^- located above k' [i.e., (k^-k') is $\perp$ to

our plane] at a distance ρ_k from it. The inequality defining $\mathcal{E}_k$ is thus transformed into

$$a_k F_k \geq s_k F_{k^-},$$

which formula corresponds in $\mathcal{R}^3$ to the volume obtained by rotating an area of the type we studied in the preceding chapters around its symmetry axis ($kk^\sim$); whence it is clear that if the area is starshaped w.r.t. k , or convex, that volume has the same property; and so has $\mathcal{E}_k$, being the intersection of that volume with our plane. Concerning item (c), we also have to notice that if the volume degenerates into a set contained in (kk''), its intersection with the plane is a point, or is empty, and is thus simply connected. *Mutatis mutandis*, all those reasonings hold for $\mathcal{E}_{k^-}$, taking Prop. 5.8 into account. Last, Fig. 9.4 makes easier to understand item (e): because of the similarity of two triangles, we have $x/\rho_k = (\rho_k/2)/\delta_{kk''}$. Q.E.D.

In particular, when $h = \ln [(\cdot)^2 + K]$ with $K \geq 0$, the volume is a sphere or the closure of the exterior of a sphere, so that $\mathcal{E}_k$ is a disk or the exterior of the closure of a disk, the centre and radius of which may easily be computed; see Prop. 2.7e.

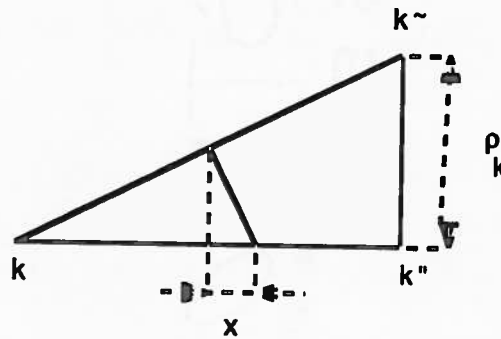


Fig. 9.4. — Proof of Prop. 9.3e.

We might also think of the area $\mathcal{E}'_k = \{i: a_k F_k \geq s_k F(\delta_{k''})\}$. On the basis of a result of Hillsman and Rhoda (1978) (namely, that $\sum_{c \in C - \{k\}} \lambda_{kc} \delta_c \geq \delta_{k''}$), it is easily shown that $\mathcal{E}'_k \subseteq Y_k$ when F is concave. However, in that case, $F \circ \sqrt{\cdot}$ is also concave and, according to Prop. 9.3a, $\mathcal{E}_k \subseteq Y_k$. As we obviously have $\mathcal{E}'_k \subseteq \mathcal{E}_k$, the area $\mathcal{E}'_k$ is a looser bound than $\mathcal{E}_k$, and thus not to be used.

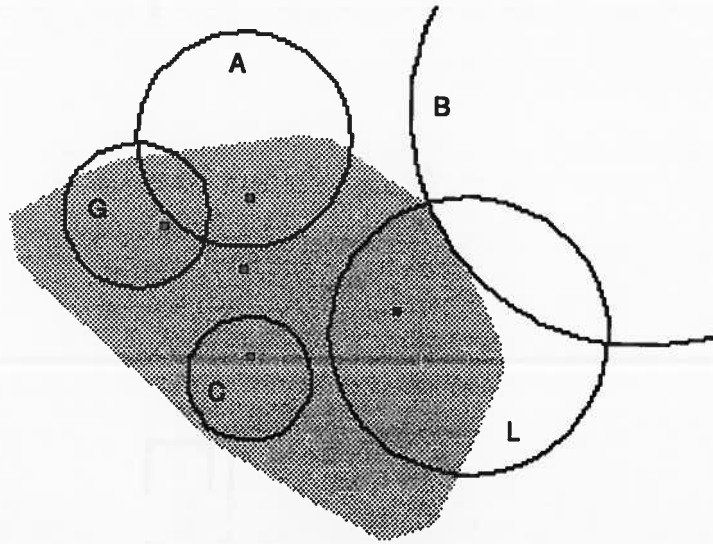


Fig. 9.5. — Approximation of dominance areas by areas $\mathcal{A}_k$; see Fig. 6.5.

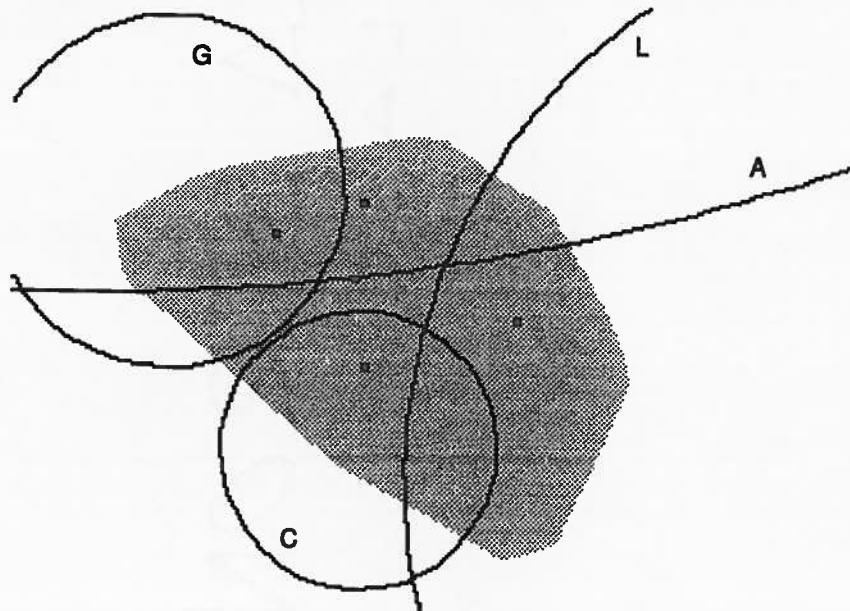


Fig. 9.6. — Approximations of arithmetic average areas by areas of type $\mathcal{A}_k$; see Fig. 6.6.

No boundary appears for Brussels.

Last, from all this, we may deduce conditions concerning the emptiness of Y_k :

(9.4) (a) When $F \circ \sqrt{\cdot}$ is convex: if h is concave and $a_k/s_k < F(\rho_k)/F_{kk}$,

Y_k is empty;

(b) When F is convex:

(i) if h is concave and if, for some point p , $a_k/s_k < F(\omega_{kp})/F_{kp}$, then Y_k is empty;

(ii) if h is convex and if, for some point p , $a_k/s_k < \exp[-h'(\infty)(\delta_{kp} + \omega_{kp})]$, then Y_k is empty;

(c) When $F \circ \sqrt{\cdot}$ is concave on the range of distances considered: if h is concave and $a_k/s_k \geq F(\rho_k)/F_{kk}$, Y_k is not empty.

9.2. Total potential

As we have emphasized in Section 1.6, speaking of total potential has a meaning in the generalized gravity model. The area $\Pi(\tau)$ where the total potential W_i , defined as

$$W_i = \sum_{c \in C} V_{ic},$$

has at least some given level τ is endowed in that model with the following obvious properties (m_{cp} is the middle of $[cp]$):

(9.5) (a) For any point p of the plane,

(i) if $p \in \Pi(\tau)$, $\bigcap_{c \in C} D_c(\delta_{cp}) \subseteq \Pi(\tau)$;

(ii) $\Pi(\tau)$ is starshaped w.r.t. p in the area

$$[\bigcup_{c \in C} D_{m_{cp}}(\delta_{cp}/2)]^-;$$

(iii) $\Pi(\tau)$ is antistarshaped w.r.t. p in the area

$$\bigcap_{c \in C} D_{m_{cp}}(\delta_{cp}/2);$$

(b) If F is concave on the range of distances considered,

$\Pi(\tau)$ is convex.

Item (a.i) simply expresses that if we decrease all the distances at the same time, W_i increases. The items (a.ii) and (a.iii) are in the same vein: on any straight line passing through p , the distance to a centre c decreases when we move toward p , except between p and the orthogonal projection of c on the line; i.e., except in the interior of the intersection of the line with the disk $D_{m_{cp}}(\delta_{cp}/2)$: hence item (a.ii). On the contrary, in that intersection, when we move toward p , the distance to c increases; which yields item (a.iii). As to item (b): the fact that $F(\delta_c)$ is concave and decreasing w.r.t. δ_c obviously implies that it is a concave function of the coordinates (x, y) of i . Being a weighted sum of such functions, W_i is also concave w.r.t. (x, y) . Hence the result. **Q.E.D.**

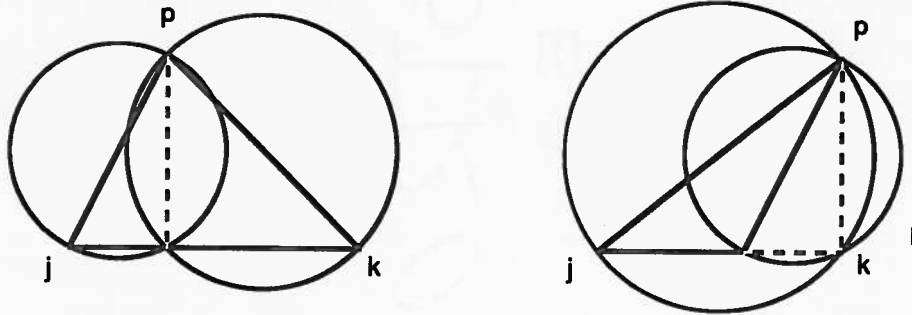


Fig. 9.7. — $D_{m_{jp}}(\delta_{jp}/2) \cup D_{m_{kp}}(\delta_{kp}/2)$ contains the triangle $[pjk]$.

If we know that a point lying outside the convex hull of C belongs to $\Pi(\tau)$, item (a) may tell us that some part of the plane is contained in $\Pi(\tau)$. But notice that when p is in the convex hull of C , the variations of the distances to the centres are necessarily conflicting. As a consequence, the area $\bigcap_{c \in C} D_c(\delta_{cp})$, concerning item a, is reduced to $\{p\}$ when p is in the convex hull. On the other hand, whatever p , the area $\bigcup_{c \in C} D_{m_{cp}}(\delta_{cp}/2)$ contains the convex hull of C ; indeed, given any two centres j and k , $D_{m_{jp}}(\delta_{jp}/2) \cup D_{m_{kp}}(\delta_{kp}/2)$ contains the triangle $[pjk]$; see Fig. 9.7. Similarly, it can be shown that the convex area $\bigcap_{c \in C} D_{m_{cp}}(\delta_{cp}/2)$ has exactly one intersection point with the convex hull; that point is the point of the convex hull closest to p (the area is thus reduced to $\{p\}$ when p belongs to the hull).

Item (b) applies in particular to the deterrence functions $\max \{0, \delta - \underline{\delta}\}$ and $\max \{0, \delta^2 - \underline{\delta}^2\}$ that we have met in Section 7.2. It is easily seen that if we may encompass the area under study, containing C , in a disk D of diameter $\underline{\delta}$, then $\Pi(\tau) \cap D$ is convex and, if not empty, contains the Weber point of C in the former case, or the inertia centre of C in the latter; the centres being of course weighted with their attractivities.

Notice that Lemma 7.1 might give some hope in further investigations of $\Pi(\tau)$, since the inequality defining $\Pi(\tau)$ may be rewritten

$$W_i = \sum_{c \in C} a_c F_c \geq \tau$$

or

$$W_i = F_k (a_k + \sum_{c \in C - \{k\}} a_c F_c / F_k) \geq \tau$$

or

$$W_i = F_k [a_k + \sum_{c \in C - \{k\}} a_c / \exp(h_c - h_k)] \geq \tau.$$

On any straight line passing through k , we have $dF_k/d\delta_k \leq 0$; if $d(h_c - h_k)/d\delta_k \geq 0$ for any centre $c \neq k$, we shall thus have $dW_i/d\delta_k \leq 0$. However, the results so obtained through the items (a) and (d) of the Lemma are weaker than Prop. 9.5b: this trail does not lead to any new proposition.

On the other hand, approximations of $\Pi(\tau)$ may be found the same way we found $\mathbb{G}_{kp}$ and $\mathbb{A}_k$ as approximations of Y_k . Let, for some point p ,

$$s = \sum_{c \in C} a_c$$

$$\lambda_c = a_c / s$$

$$\omega_p = \sum_{c \in C} \lambda_c \delta_{pc}$$

$$\Pi'_p(\tau) = \{i: F(\delta_p + \omega_p) \geq \tau / s\}$$

and

$$\rho^2 = \sum_{c \in C} \lambda_c \delta_c^2 - \delta_g^2,$$

which is, as we have seen in Section 9.1, independent of i ; and let

$$\Pi''(\tau) = \{i: F[(\delta_g^2 + \rho^2)^{1/2}] \geq \tau / s\}.$$

Both approximate areas are disks, centred on p or g , respectively. Similarly to Prop. 9.1 and 9.3, we have:

- (9.6) (a) If F is convex, $\Pi'_p(\tau) \subseteq \Pi(\tau)$;
- (b) (i) If $F\sqrt{\cdot}$ is convex, $\Pi''(\tau) \subseteq \Pi(\tau)$;
- (ii) if $F\sqrt{\cdot}$ is concave on the range of distances considered,
 $\Pi(\tau) \subseteq \Pi''(\tau)$.

Of course, here again, the first bound is in fact valid whatever the transportation space, whereas the second is directly related with our particular transportation space. And the Weber point, w , of the set C , every centre c being assigned a weight a_c , seems a good choice for p ; the reason is very clear when F is convex, since this choice maximizes the radius of $\Pi'_p(\tau)$ and thus the measure of this lower bound on $\Pi(\tau)$.

The best conditions that may be deduced from Prop. 9.6 concerning the emptiness of $\Pi(\tau)$ are the following:

- (9.7) (a) If F is convex and if $F(\omega_w) \geq \tau / s$, $\Pi(\tau) \neq \emptyset$;
- (b) if $F\sqrt{\cdot}$ is concave on the range of distances considered
 and if $F(\rho) < \tau / s$, $\Pi(\tau) = \emptyset$.

The fact that ω_w is the best ω_p in item (a) is obvious since it minimizes ω_p ; moreover, $\omega_w \leq \rho$ since $\omega_g \leq \rho$, as ω_g is the average of the distances from the centres to g and ρ is the square root of the average of the squares of those distances. The condition ' $F(\rho) \geq \tau/s$ ' that might be deduced from Prop. 9.6b.i would thus be weaker than Prop. 9.7a.

10. Confrontation with reality...

10.1. Zero-dimensional areas

We shall extend here, but not explore thoroughly, the reflexions of Section 5.7 (Fig. 5.4 and 5.6). Is it possible to deduce 2-dimensional areas from 0-dimensional ones like that of Fig. 10.1 ? Otherwise said: given some point i of Y_k or Z_k , is it possible to deduce a 2-dimensional area contained in that set ? Obviously, we remain in the area if we decrease δ_k while increasing the distances to the other centres; thus for each point i of the area, $D(k, i) = \cup \{D^0(c, i) : c \in C\}$ is contained in the area. However, that set has an empty interior when k is in the convex hull of the other centres; the property is useful only when k is one of the corners of the convex hull of C .

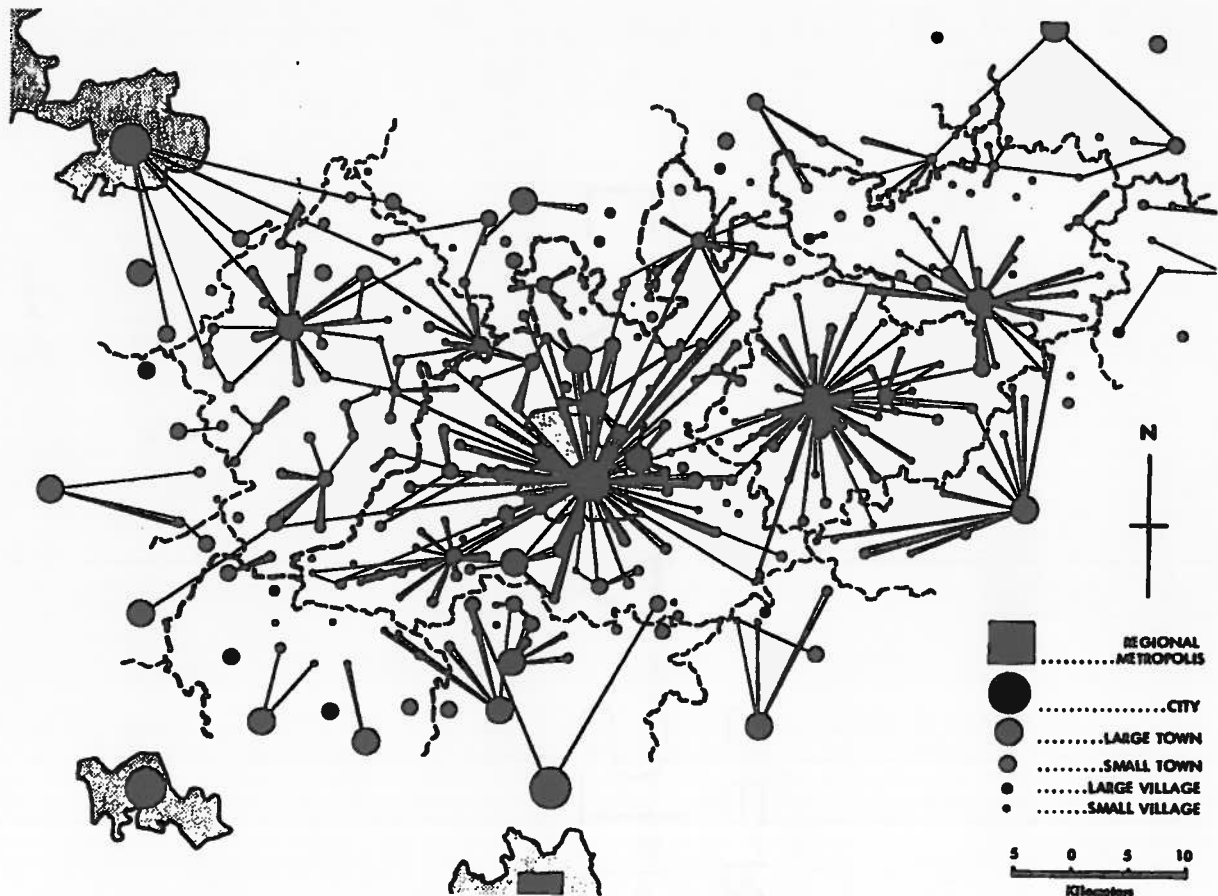


Fig. 10.1. — Market areas for movie and theatre entertainment in Baden-Württemberg; excerpt from Gardiner Barnum (1966), fig. 27.

More can be said, however, about market areas. First, it is possible to study the ranking of the centres w.r.t. attractiveness, relying on the sets

$Z_c \cap M$. Indeed, if a point i belongs to Z_k but is farther from k than from some other centre, say j (i.e., $\delta_{ik} > \delta_{ij}$), then $k \succ j$. We thus see that the centres contained in $D^0(i, k)$ are less attractive than k . On this basis, we may already decide to stop the study if too many contradictions occur, or to go on with it after eliminating erratic data, if any.

As to the 2-dimensional area B_{ik} containing i and $\subseteq Z_k$, we determine it as $\bigcap \{B_{ikc} : c \in C - \{k\}\}$. If a centre j is contained in $D^0(i, k)$, we know that B_{ikj} is as depicted on the right part of Fig. 5.4; it may be increased if we know more about the function, as in Fig. 5.6.

But what if j is not contained in $D(i, k)$? If we know that $k \succ j$, B_{ikj} is here simply the half-plane (independent of i) containing k and the boundary of which is the mediatrix of $\{j, k\}$. If we do not know whether k is more attractive than j , we may as well do as if k were less attractive than j , since the area B_{ikj} we shall compute will be contained in that half-plane: we shall not add undue points to B_{ikj} . As to that B_{ikj} , it is easy to draw if we give ourselves a precise t.c.f.: B_{ikj} is the area Z_{kj} corresponding to values of the attractivenesses a_j and a_k for which $Z_{kj} \cap Z_{jk}$ passes through the point i .

If we do not assume a precise t.c.f., B_{ikj} is then at least as in the left part of Fig. 5.4, and may again be enlarged if we know more about the t.c.f. If for instance the t.c.f. is supposed to be concave, B_{ikj} is to be drawn as in Fig. 5.6. Of course, if the t.c.f. is supposed to be convexifying, we have also to consider all the points i' for which $j \notin D(i', k)$: the convex hull of the union of the areas $B_{i'kj}$ is contained in Z_k (possibly, new points i'' appear for which $B_{i''kj}$ is not enclosed in the area retained so far; a new convex hull is then to be drawn, and so on).

10.2. Market and shared influence areas

Some maps display, between areas that are in fact *strict* market areas, intermediate areas where it is said that the influence of no centre is dominant. This is of course not strictly true. Indeed, those intermediate (or shared influence) areas concern situations in which the difference between the influences of the two most attractive centres is negligible; i.e., smaller

than some given threshold. As in Section 1.7, we define extended market areas by decreasing slightly p_k (i.e., increasing slightly a_k), and the shared influence areas are created by the intersection of extended market areas. We may so write the extended market area of k as $Z_k[a_k(\eta)]$, for some η slightly > 1 (see Section 6.2).

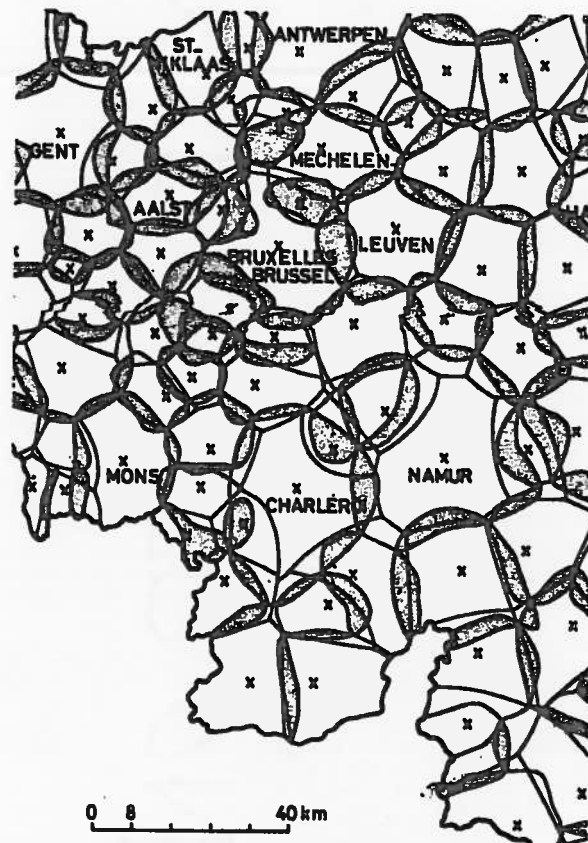


Fig. 10.2. — Excerpt from the map of the first level of urban hierarchy in Belgium; from Beguin and Thill (1985).

The map 28 of the *Atlas of Belgium* (1950-1972) displays influence areas ('zones d'influence') of Belgian cities; see Fig. 10.2, 10.4, and 10.5 ¹. The areas overlap, letting appear shared influence areas. The comparison beared in fact on several goods. A commune was included in the strict market area of a city if its buyings were 'essentially' performed from that centre; whereas it was included in a shared influence area when a discrepancy appeared

¹ Other works containing such economic maps are listed by Thill (1982, 1983) and Beguin and Thill (1985).

between the buying places of some characteristic goods, or between the economic polarization of hamlets. Moreover, the maps so obtained were simplified: curves were smoothed, and extended market areas were made simply connected; see Thill (1982). We are thus rather far from our model. Nevertheless, it seems reasonable to think, together with Thill, that we face here something close to a set of maps of extended market areas.

Some facts, however, are surprising in that perspective. The main of them appears in the map related to the first level of urban hierarchy; see Fig. 10.2. Most of the extended market areas have indeed a rounded shape; i.e., a shape that looks nearly convex. This seems rather strange if we accept the assumption that most centres have roughly the same attractiveness at this level.

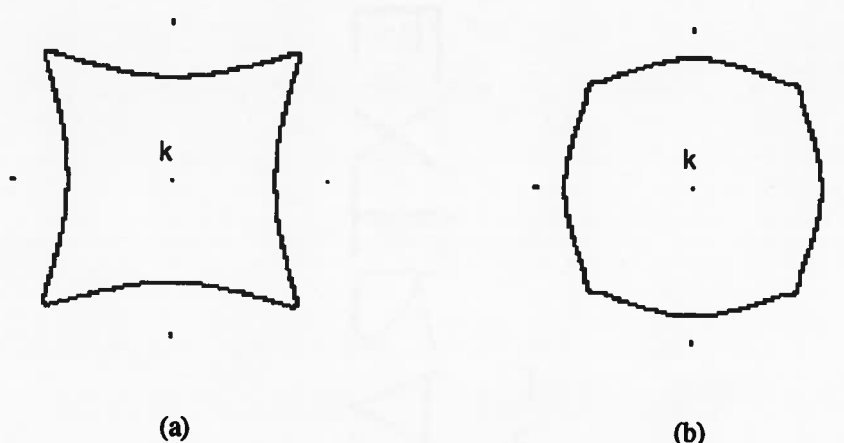


Fig. 10.3. — Theoretical extended market areas. In both cases, the shortest distance between two centres is 3; all attractivenesses are equal.

On the left, $F(\delta) = \exp(-0.1 \delta)$ and $\eta = 1.1$;

on the right, $F(\delta) = \exp(-0.005 \delta^4)$ and $\eta = 1.2$.

Consider indeed two equally attractive centres, j and k ($Q = 0$). It is clear from Prop. 3.12 that the difference in abscissa between the boundaries of the extended market areas $Z_j[\underline{a}_j(\eta)]$ and $Z_k[\underline{a}_k(\eta)]$ is an increasing function of the ordinate when the function $h\sqrt{\cdot}$ is concave, and a decreasing when $h\sqrt{\cdot}$ is convex. More simply said: starting from their intersection with the line (jk) , the two curves diverge in the former case and converge in the latter. This has the consequence that $Z_j[\underline{a}_j(\eta)]$ and $Z_k[\underline{a}_k(\eta)]$ have a kind of round shape when $h\sqrt{\cdot}$ is strictly convex; for instance when $h(\delta) = \delta^a$ with $a > 2$. The round shape of the first-level extended areas in the *Atlas* thus suggests

that property. Similarly, the fact that several centres do not belong to their strict market area is incompatible with the concavity of the t.c.f. As we have said in Section 2.5, one might expect the t.c.f. to be concave.

In opposition to the preceding observations, in the map of the third level (Fig. 10.4), the extended market area of Brussels is absolutely not starshaped by that city, due mainly to the shape and position of the strict market area of Gent. Brussels being the most attractive centre, the t.c.f. can thus not be convex. Even the linear t.c.f. should here be discarded.

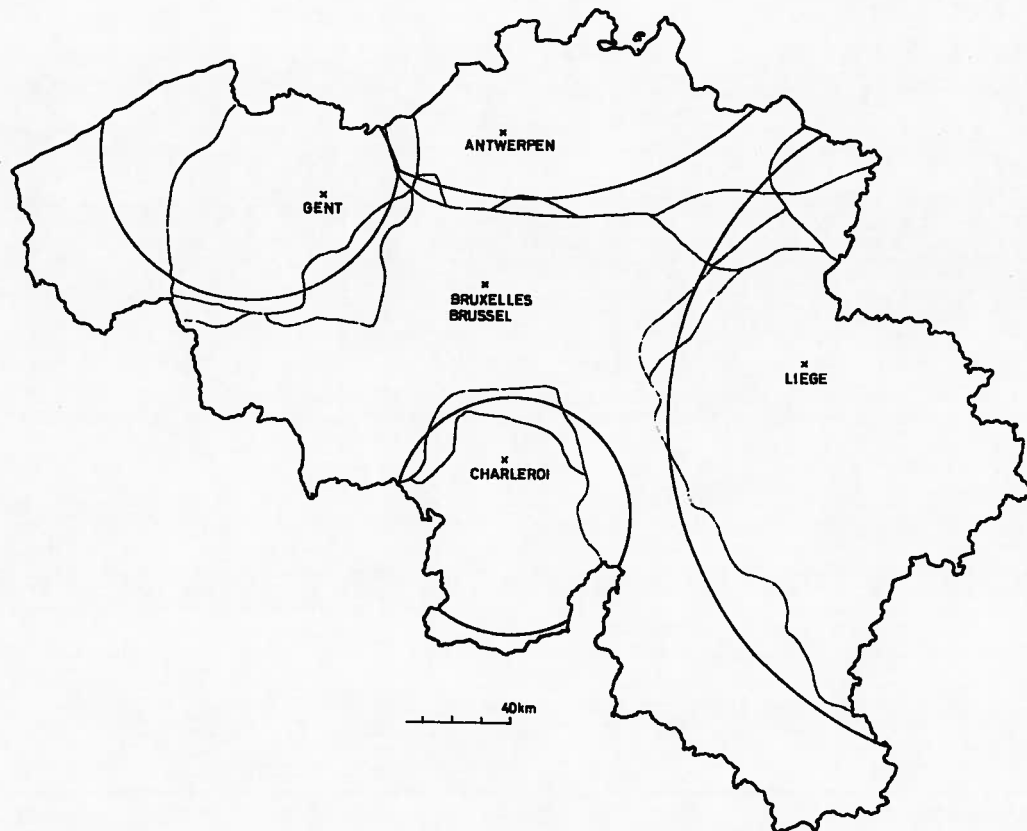


Fig. 10.4. — Map of the third level of urban hierarchy in Belgium;
from Beguin and Thill (1985).

The contradiction between our conclusions concerning the levels I and III emphasizes the importance of treating each level separately. Thill did so and computed that the exponent a of his Pareto model was equal, respectively, to 6.0, 7.5, and 3.2 in the levels I, II, and III. But the t.c.f.'s remain proportional to $\ln(\delta)$ in those three cases, so that nothing is fundamentally changed in the shape of theoretical market areas. Perhaps more radically

different t.c.f.'s should be tried. However, at the first level, the fact that the extended market area of Namur has no round shape is absurd if we admit that $h \cdot \sqrt{\cdot}$ is convex, since Namur is locally most influent. The best conclusion seems that the generalized gravity model is not suitable for the first level.

Notice also that the common boundary between two market areas must turn its concavity towards the less attractive centre when h is concave, as we have seen concerning the two-centre problem at the end of Section 4.7. In one case at least, at the second level, the shape of the observed common boundary contradicts that of the theoretical one as computed by Thill with $\ln(\delta)$ as t.c.f. At the second level of urban hierarchy, it seems indeed clear that the real indifference line between Hasselt and Leuven turns its concavity towards Leuven. The real attractiveness of Hasselt should thus be larger than that of Leuven. In opposition to this, however, the breaking point on the straight segment joining those two cities is closer to Hasselt.

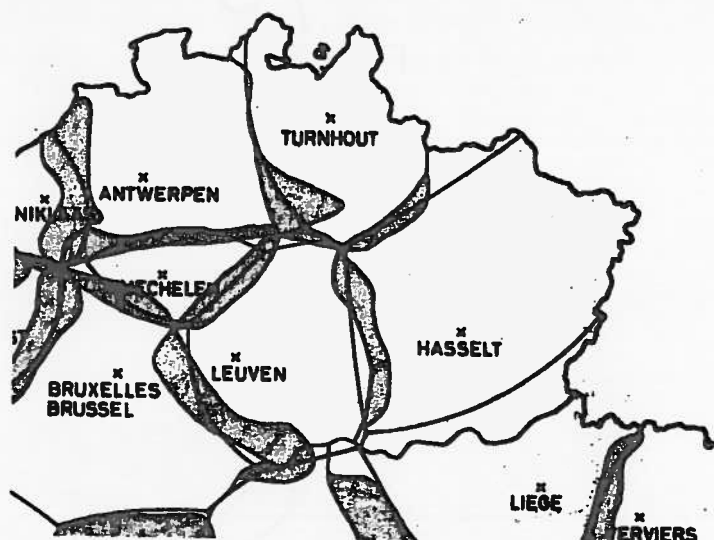


Fig. 10.5. — Excerpt from the map of the second level of urban hierarchy in Belgium [from Beguin and Thill (1985)]

10.3. Share areas

Other interesting maps are those realized by Barbier (1969) and Bruyelle (1970): we have here maps covered with τ -share areas for several values of τ . Fig. 10.6 shows such areas for τ equal to 40% and 10% in the vicinity of Boulogne according to the latter author. The 10%-share area of

Boulogne is absolutely not starshaped w.r.t. Boulogne; the 40%-share area is also not perfectly starshaped w.r.t. that city. We can at least draw the conclusion that the t.c.f. is not linear. But, even if we only assume h to be concave, the 10%-share area should be starshaped w.r.t. Boulogne since it is contained in the Voronoi cell of that centre. Maybe the area *is* starshaped w.r.t. Boulogne, however, but we cannot see it because our approximation of the transportation space as Euclidean is too rough; remember that those properties of starshapedness of market and dominance areas related with the concavity, convexity, or linearity of the t.c.f. are valid in any transportation space.



Fig. 10.6. – Share areas at the levels 10% and 40% of the metropolis and main centres for domestic electrical appliances, radio sets and televisions in the region of Lille (France); excerpt from Bruyelle (1970), plate 4.1.

Conclusion

Let us shortly summarize our study and its results. The main object of our investigation is the gravity model. This model is the separable multiplicative mathematical expression of a very simple idea: the level of interaction between two entities is an increasing function of their respective importance, and a decreasing one of the distance between them. The interaction t_{ic} between a consumer location i and a centre c is thus here represented by the quantity $m_i a_c F(\delta_{ic})$; m_i is the mass of the consumer location, a_c is the attractiveness of the centre, and F is the distance deterrence function.

Our work also concerns less ambitious models in which the interaction of a consumer location with several centres is only assumed to be distributed among them proportionally to quantities $a_c F(\delta_{ic})$. One aspect of the gravity model is thus kept here, and the other is dropped; namely, that the interaction with a centre c should also be distributed among the consumer locations proportionally to the quantities $m_i F(\delta_{ic})$. Several models fit that model or the gravity one, as far as the interactions t_{ic} are concerned; among them are models trying to give a theoretical basis to the gravity model.

The first part of our study explores properties of market areas when there are only two centres, j and k . But what is a market area? Defining the area of k as the area where $t_{ik} \geq t_{ij}$ has the consequence that the consumer places where $m_i = 0$ are contained in both market areas. The market area of k , Z_k , is thus better defined independently of the masses by the inequality $V_{ik} \geq V_{ij}$, in which the quantity V_{ic} , known as a potential, is equal to $a_c F(\delta_{ic})$.

As far as market areas are concerned, the models we are investigating are formally equivalent, through a logarithmic transformation [$h = -\ln F$, $p_c = -\ln(a_c)$], to rational models in which the consumer located in i chooses the centre he will interact with by comparing the total prices $p_c + h(\delta_{ic})$ (mill prices plus unit transportation costs) of interacting with the centres. As a matter of fact, that additive formulation is found to be particularly convenient for our study, so that we express the problem in terms of transportation cost functions in any case. It is worth noticing that, if we agree that transportation costs should be a concave function of distance, function h is likely to be concave since it appears in many models as proportional to transportation costs or to their logarithm.

Conclusion.2.

Let us first have a look at the simple cases listed at the end of Chapter 2. With a distance deterrence function proportional to $\delta^{-\beta}$ with $\beta > 0$ (i.e., in the Pareto model), the market area of k (supposed to be less attractive than j) is a disk. When $F(\delta)$ is proportional to $\exp(-\beta\delta)$ (the well-known negative exponential model), Z_k is a convex area limited by one of the two branches of a hyperbola. When $F(\delta)$ is proportional to $\exp(-\beta\delta^2)$ (normal or Gaussian model), Z_k is a half-plane. When it is proportional to $(\delta+K)^{-\beta}$ with $K > 0$, Z_k is a convex set limited by a Cartesian oval; last, when $F(\delta)$ is proportional to $(\delta^2 + K)^{-\beta}$ with $K > 0$, Z_k is again a disk. Only one disturbing phenomenon is revealed here: the centre k may belong to the market area of the other centre, Z_j , when $F(\delta)$ is proportional to $\exp(-\beta\delta^2)$ or to $(\delta^2 + K)^{-\beta}$. Apart from that, the impression is that market areas have very gentle shapes.

However, the basic condition that h should be strictly increasing (or F strictly decreasing) is not sufficient to guarantee such simple shapes. A striking instance of this appears in Fig. 2.2, in which we use the t.c.f. $\delta^2 + \sin(\delta^2)$. Another interesting example is given by Fig. 5.1a: when the attractiveness constant Q (i.e., the difference $p_k - p_j$) is 0, the two market areas are half-planes; if Q increases and exceeds a threshold value approximately equal to 1.458, a finite arcwise connected component of Z_k emerges; it vanishes for higher values of Q . Rather irregular shapes may appear even if we require in addition h to be concave, as shown for instance by Fig. 5.3a, where $h(\delta)$ is the lower envelope of three linear functions.

We shall not treat here of Chapter 3, which mainly contains intermediate results. The general purpose of Chapter 4 is to find some bounds on market areas. In particular, it shows that j , the more attractive centre, always belongs to Z_j , and that every nonempty market area contains its centre when h is concave; those two results are valid in any transportation space (even inside a network, for instance). It also appears that Z_j is unbounded; and that, when h is convex, Z_k , if not empty, is also necessarily unbounded. Through a study of circular bounds on Z_k , a salient characteristic is also detected: with some transportation cost functions, the market area Z_k may be more extended in the direction perpendicular to the straight line (jk) than along that line; see Corollary 4.1 and Fig. 4.3 and 5.1b. That property holds, in particular, for the t.c.f. $-\delta^{-a}$, with $a > 0$; whereas, when the t.c.f. is for instance δ^a with $a \in]0, 1[$ or $\ln(\delta + K)$ with $K > 0$, Z_k is more extended in the direction of (jk) .

Conclusion.3.

Still in Chapter 4, we also examine some simple unbounded curves encompassing Z_k or contained in it; this leads us to study the asymptotic properties of the indifference zone $Z_j \cap Z_k$. In particular, we so detect some peculiar possible shapes of Z_k depicted in Fig. 4.5 and 4.6. The distance deterrence function $\exp(-\beta\delta)/\delta^{\beta'}$ used by Tanner and others, intermediate between Pareto and negative exponential models, is considered in that perspective in Fig. 4.2 when $\beta = \beta'$.

Chapter 4 tries to give a global idea of the market areas, as if seen from afar; in opposition to this, Chapter 5 is concerned with their internal structure. In the latter chapter, we first examine if the market areas are arcwise connected; i.e., if it is possible to go from one point of the area to another without leaving the area. The main results are that Z_j is always arcwise connected, and that, if h is concave or convex, Z_k must also be arcwise connected. On the other hand, Z_k is always simply connected, which means that it has no hole. As to Z_j , it is simply connected when h is convex; but not when h is concave, unless Z_k is unbounded. Another property is that Z_j is spined, as we say, by the y -axis (the line perpendicular to the segment $[jk]$ in its middle) when h is convex; i.e., the segment joining any point of Z_j to its orthogonal projection on the y -axis is in Z_j . It appears that Z_j is also spined by the line (jk) when the function $h \circ \sqrt{\cdot}$ is convex; and that Z_k is spined by (jk) when $h \circ \sqrt{\cdot}$ is concave.

Two major properties are also found in Chapter 5. The first holds in any transportation space: if h is concave, Z_k is starshaped by k (i.e., the shortest path between k and any point of Z_k is a part of Z_k); if h is convex, Z_j is starshaped by j ; and if h is linear, each market area is starshaped by its centre. The second property identifies some transportation cost functions for which Z_k is convex in our particular transportation space; among them are the functions $-\delta^{-a}$ with $a > 0$, δ^a with $a \in]0, 1]$, and $\ln(\delta + K)$ with $K \geq 0$. As a preliminary to Part II, the last two sections of Chapter 5 discuss some problems concerning topological concepts like boundaries and closures, and examine how we can deduce from the knowledge of isolated points of Z_k or Z_j that some two-dimensional regions are contained in those areas.

The second part of this work is dedicated to market areas when there are more than two centres. However, we first have to reconsider the basic question: what is a market area? The definition used in Part I may indeed be extended in several directions. The area of a centre, k for instance, may still be the area in which the potential V_{ik} is not smaller than all the other

Conclusion.4.

potentials; in which case we keep the name of market area. But V_{ik} may also be compared with the sum of all other potentials; the area, that we now call the dominance area of k , is then the region where V_{ik} is not less than 50% of the total potential. A third possible area is where V_{ik} is not less than some average of the other potentials, or than their median; of course, an infinity of averages are possible. Other definitions found in the literature are also presented; as well as related subjects: choice, share, and preference areas. That variety has its origin in the fact that we try to give a very condensed idea, through a set of n curves drawn on the plane (n being here the number of centres), of the information yielded by the n -dimensional vectors (V_{ic} : $c \in C$) attached to each point of the plane.

In Chapters 7 and 8, we see that the properties of starshapedness and convexity discussed in Chapter 5 are easily extended, in the many-centre problem, to market areas and, as far as starshapedness is concerned, dominance areas; other areas are also briefly discussed. Concerning linear transportation cost functions (i.e., the negative exponential model), in particular, a strong property is found: all market and dominance areas are starshaped by their centre (Fig. 7.7 and 7.8); with concave t.c.f.'s, the least attractive centre starshapes its market area, whereas, with convex t.c.f.'s, the most attractive centre displays that feature. Here again, those three properties hold in any transportation space. Starshapedness is thus important from both the theoretical and practical points of view. Theoretical: because it is not infrequent to meet in the literature the implicit assumption that market areas, or others of a similar type, are starshaped w.r.t. their economic centre; as in Huff (1963), writing that what we have called share areas radiate from their centre (at least for the less attractive one), Huff and Batsell (1977), Huff and Rust (1984), or Goodchild (1987). Practical: because it may lessen the number of computations required by some GIS computer programs, like that of Miller (1994).

Still concerning starshapedness, the study of the two-centre problem is deepened in Chapter 7 in order to get better results than those directly deducible from Part I. With the Pareto model, for instance (i.e., when h is logarithmic), by comparison with what is deducible from Part I, we so multiply by four the superficies of the region inside which we know *a priori* that the market and dominance areas of a centre must be starshaped by that centre; see Fig. 7.5, in which the property appears very accurate: the boundary of the dominance area, if intersected by a wider region, is no more entirely visible from the corresponding centre. On the other hand, the regular organization of

Conclusion.5.

some patterns of centres, like those of central place models considered by Parr (1995b), may yield some results; see Fig. 7.6. Also, in Chapter 7, we are led to consider distance deterrence functions having a finite range, because two such functions have the remarkable property that dominant centres starshape their dominance areas (for one of them, the property is independent of the transportation space); see Fig. 7.9 and 7.10.

We also discuss in Chapter 7 whether a centre belongs to its market or dominance area; let us point out these results: if h is concave, each market area (if not empty) contains its centre; if $h(0) = -\infty$, each market or dominance area contains its centre; the most attractive centre always belongs to its market area; dominant centres (i.e., gathering at least 50% of the total attractiveness) belong to their market and dominance areas. Another aspect of Chapters 7 and 8 is that we try to give a precise meaning to the intuitive idea that the market area of a centre frequently depends only on its 'neighbour' centres.

One reason for which starshapedness and convexity appear as important properties is that they can be easily studied when there are more than two centres. Starshapedness is obviously more important since it generally concerns all transportation spaces, as we have seen. However, the interest of convexity is stressed in Chapter 8. The study of the three-centre problem leads us indeed to identify some properties, the main of which is that, if all centres are on a same straight line, all market areas are arcwise connected provided that the t.c.f. is concave and strictly convexifying [i.e., yielding a strictly convex area Z_k in the two-centre problem, unless Z_k is contained in the straight line (jk)]. That condition is satisfied, in particular, by the transportation cost functions $-\delta^{-a}$ (with $a > 0$), $\ln(\delta)$, δ^a (with $a \in]0, 1[$) and $\ln(K + \delta)$ (with $K > 0$). Those statements might be extended to other transportation spaces if we could detect transportation cost functions having the same properties in those spaces.

Chapter 9 gives bounds on dominance areas and total potential. The dominance area of a centre, in particular, is approximated by the market area the centre would have if it faced only one other centre gathering all the remaining attractiveness and located some distance from the plane, being either related through a path to the Weber point of the other centres, or situated above their gravity centre (the mass allocated to each centre being equal to its attractiveness).

Conclusion.6.

The last chapter, Chapter 10, is dedicated to applications. We first develop the method, already sketched in Chapter 5, to be followed in order to derive two-dimensional market areas from the knowledge of zero-dimensional ones; i.e., of sets of consumer points allocated to one centre or another. We then review some empirical economic maps of market and dominance areas and see if they fit our theoretical results. Maybe the most interesting case appears in the map of market areas of the cities of the first level of the urban hierarchy of Belgium. Those areas, which, as a matter of fact, are extended market areas (i.e., including shared influence regions), tend indeed to have a nearly convex shape; a phenomenon indicating that the t.c.f. is very convex, more than δ^2 . Other features, however, contradict that conclusion and suggest a concave t.c.f.

Some other questions have been met at several moments of this research. How far can we trust the generalized gravity model? Is the Pareto model better than the negative exponential one? Although we did not make a systematic investigation of those issues, I hope we have clarified them a little. Similarly, I am glad to have shown that choice and share areas could be included in our reflection. Something, maybe, is a bit disappointing: some intuitively obvious properties were so difficult to demonstrate! This is an indication that either (i) we have to find better proofs, (ii) we should not put too much confidence in intuition, or (iii) intuition is much more efficient than mathematics...

والله اعلم - God knows better.

Appendix : a program written with Mathematica

```
- Model data : attractivenesses ('attr') and coordinates ('v') of the five
centres
attr={0.353332, 1., 0.116996, 0.141004, 0.24636}
v={{8.11,13.125},{7.83,10.45},{8.21,7.34},{5,12},{13.45,9.07}}

- Basic calculations : distances to centres ('dist'), distance deterrence
function ('f'), partial potentials ('partpot')
d[x_,y_,i_]:=Sqrt[(v[[i]][[1]]-x)^2+(v[[i]][[2]]-y)^2]
f[x_,y_,i_]:=1./d[x,y,i]^3.2
partpot[x_,y_,i_]:=f[x,y,i]*attr[[i]]

- Data concerning the five centres and the rough limits of Belgium
gg=Graphics[{GrayLevel[0.9],Polygon[{{-0.05,10.27},{-0.23,10.55},
{-0.48,12.33},{2.05,13.9},{3.32,14.33},{9.75,15.33},
{10.99,15.225},{14.66,12.875},{16.73,9},{17.28,7.77},
{17.43,6.82},{15.02,1.32},{14.81,1.35},{13.29,.8},
{13.05,.75},{7.125,4}}]}]
fr=Graphics[{GrayLevel[0.1],PointSize[0.01],
Map[Point,Table[v[[i]],{i,5}]}]}]

- Market areas
listek[x_,y_,k_]:=Delete[Table[partpot[x,y,i],{i,5}],k]
pref[x_,y_,k_]:=partpot[x,y,k]-Max[listek[x,y,k]]
grpref[k_]:=ContourPlot[pref[x,y,k],{x,-15,25},
{y,-10,25},ContourShading->False,PlotPoints->40,
Contours->{0.}]
Show[gg,fr,grch[1,3],grch[2,3],grch[3,3],grch[4,3],grch[5,3],
AspectRatio->Automatic]

- Choice areas:  $Ch_k(m+1)$ 
listemod[x_,y_,k_,m_]:=Drop[Sort[listek[x,y,k]],-m]
choix[x_,y_,k_,m_]:=partpot[x,y,k]-Max[listemod[x,y,k,m]]
grch[i_,k_]:=ContourPlot[choix[x,y,i,k],{x,-15,25},
{y,-10,25},ContourShading->False,PlotPoints->15,
Contours->{0.}]
Show[gg,fr,grch[1,3],grch[2,3],grch[3,3],grch[4,3],grch[5,3],
AspectRatio->Automatic]
```