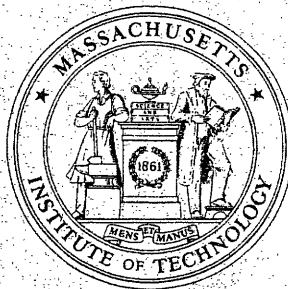


# OPERATIONS RESEARCH CENTER

working paper



**MASSACHUSETTS INSTITUTE  
OF TECHNOLOGY**

LARGE SCALE GEOMETRIC LOCATION PROBLEMS I:  
DETERMINISTIC LIMIT PROPERTIES

by

Mordecai Haimovich\*

OR 130-85

Revised October 1985

\* University of Chicago, Graduate School of Business



## Abstract

Consider a finite positive Borel measure,  $w$ , with bounded support and a set  $C$  of  $K$  points in  $R^d$ . The cost of serving a customer population distributed according to  $w$  by facilities located at the points of  $C$  is sometimes proportional to  $\int \min_{c \in C} \|x - c\| dw(x)$ . The  $K$ -median problem is the problem of finding the lowest attainable cost for a given  $w$  and the corresponding (optimal)  $C$ . We show that as  $K \rightarrow \infty$ , the optimal cost is asymptotically proportional to  $K^{-1/d}$ . Let  $\mu$  and  $m$  denote, respectively, Lebesgue's measure and the density of the absolutely continuous part of  $w$ . We further show that the proportionality constant is equal to  $\gamma_d (\int m^{d/(d+1)} d\mu)^{(d+1)/d}$ , where  $\gamma_d > 0$  is a constant independent of  $w$ , and that the density of the optimal  $C$  points at any given location is proportional to  $m^{d/(d+1)}$ .

Our proofs are general enough to accommodate several versions of the  $K$ -median problem. We also outline various extensions and modifications of the results to problems with nonlinear costs, constrained facilities and certain classes of measures  $w$ , that correspond to highly clustered customer populations.

### Outline:

1. Introduction
2. Preliminary Discussion
3. Main Convergence Results
4. Restricted Medians
5. Extensions and Modifications
6. Fractal Models for Clustered Demand
7. Summary
8. References

Appendices A-D

### Acknowledgements:

This research was supported by grants ECS-7926625 and ECS-8316224 from the Systems Theory and Operations Research Division of the National Science Foundation.

The author wishes to thank Professor Thomas L. Magnanti for his encouragement and for many helpful suggestions.



1. INTRODUCTION

Let  $x = (x_1, x_2, \dots, x_n)$  be a set of points in the plane corresponding to the locations of customers ("sinks") requiring some service (commodity). A well-known problem of spatial economics is to select locations  $c_1, c_2, \dots, c_K$  for  $K$  facilities ("sources") so as to minimize the cost  $\sum_i \min_{1 \leq j \leq K} \|x_i - c_j\|$ , where  $\|x_i - c_j\|$  denotes the Euclidean distance between  $x_i$  and  $c_j$ . This objective assumes that each customer is served by the facility nearest to it and that the Euclidean distances measure service costs. The problem is known as the Multi Source Weber Problem ([We], [K1]) and for  $K = 1$  it dates back to Fermat and Steiner ([K2], [Co]).

It is sometimes convenient, however, to consider an infinite number, or even a continuum, of customers. This type of model arises when approximating large customer populations or when modeling a random demand by a non-discrete probability distribution. In general, then, let  $w$  be a finite positive Borel measure on  $R^d$ . We refer to  $w$  as the demand or the demand distribution and assume that it is defined on some bounded region  $R$ . Let  $C$  be a finite non-empty set of points in  $R^d$  called centers that are locations for the facilities. The cost associated with satisfying the demand  $w$  by facilities located at centers in  $C$  is given by

$$D(w,C) = \int \min_{c \in C} \|x - c\| dw(x) . \quad (1.1)$$

Let  $|C|$  denote the cardinality of the set  $C$ . Then the problem of locating  $K$  centers so as to minimize cost becomes

$$\begin{aligned} & \text{Minimize } D(w, C) . \\ & \text{subject to: } |C| \leq K \end{aligned} \tag{1.2}$$

Let  $D_K(w)$  denote the optimal objective value of this problem. Borrowing terminology from graph location problems ([H2]), we refer to problem (1.2) as the (free<sup>1</sup>)  $K$ -median problem. In this paper, we study the asymptotic properties of the value function and the solution of the  $K$ -median problem (1.2) as  $K \rightarrow \infty$ . With  $m$  denoting the density of the absolutely continuous part of the demand  $w$ , and with  $\mu$  denoting Lebesgue (area) measure, we show in Sections 2 and 3 that

$$\lim_{K \rightarrow \infty} K^{1/d} D_K(w) = \gamma_d \left( \int m^{d/(d+1)} d\mu \right)^{(d+1)/d} \tag{1.3a}$$

for some constant  $\gamma_d > 0$ .

Furthermore, letting  $k_K(T)$  denote the number of centers, out of a total of  $K$ , that are located in the region<sup>2</sup>  $T$ , we show that

$$\lim_{K \rightarrow \infty} \frac{k_K(T)}{K} = \frac{\int_T m^{d/(d+1)} d\mu}{\int m^{d/(d+1)} d\mu} \tag{1.3b}$$

---

<sup>1</sup>In the usual  $K$ -median problem defined on a finite graph with nodes modeling demand locations, the centers may be placed only at nodes. In our setting, the analogous (restricted)  $K$ -median problem requires  $C$  to lie in the support of  $w$ . Observe, however, that the restricted and free versions of the problem coincide whenever the support of  $w$  is convex. We also show in Appendix C that the asymptotic behavior of these two versions is identical.

<sup>2</sup>Generally, we use the term region, without qualification, to mean a Lebesgue measurable set.

For  $d = 2$  and the special case where the demand  $w$  is uniform over some bounded region, other papers ([HM], [Pa], [F2] show that the optimal center configuration induces, asymptotically, a partition of the region into  $K$  congruent regular hexagons (as in a honeycomb) each served by a single center. Consequently,  $\gamma_2 = \phi(6)$  where  $\phi(6) = \sqrt{2/(3\sqrt{3})} (1/3 + 1/4 \ln 3)$  is the objective value of the 1-median problem when a unit of demand is distributed uniformly over a regular hexagon of unit area.

The asymptotic formulas derived in this paper are useful for establishing preliminary designs of large-scale facility location systems. In particular, they suggest an asymptotically optimal heuristic that first allocates centers (facilities) according to formula (1.3b) and then (for  $d = 2$ ) utilizes the hexagonal partitioning property to solve in small subregions where the demand is approximately uniform. A stochastic counterpart of this paper [Ha] gives a rigorous justification to the approximation of large customer populations by essentially smooth (absolutely continuous) demand distributions. Assuming that there are  $N$  customers, independently located at random, according to some probability measure  $p$ , that paper shows that if  $N/K \rightarrow \infty$ , then the cost and the solution of the  $K$ -median problem coincide (asymptotically) with the cost and the solution of an underlying deterministic  $K$ -median problem with the probability measure  $p$  as demand.

Since the validity of some of the asymptotic properties of the Euclidean  $K$ -median problem, and the techniques employed in their proof, extend to a seemingly wide spectrum of problems, we strive for generality in many of our proofs. Abstracting from the problem just the structure that is essential to each particular property, leads, we believe, to



greater insight into the nature of results and has the advantage that various extensions become more transparent.

In Section 2 we outline the central idea for our proof and then go on in Section 3 to formally establish the convergence result (1.3a, b). By abstracting the properties of the K-median problem that are essential to these results, we are able later, in Sections 5 and 6, to point out immediate extensions and modifications of the results to a variety of location problems, namely problems (i) with costs that are not necessarily linear in the customer to facility distances, or that are not necessarily linear in the demand (i.e. in the number of customers), (ii) with distances that are non-Euclidean, (iii) with capacitated facilities or with facility set-up costs, or (iv) where the demand distribution is clustered in a hierarchical fashion, as in the so-called "Fractal" (see [Ma]) model.

## 2. PRELIMINARY DISCUSSION

In this section, we informally motivate the main results that are fully established in Section 3. This preliminary discussion introduces several of the ideas that are central to the proofs, which are by necessity somewhat technically involved in places.

First, consider situations in which the demand is distributed uniformly on a (d-) cube,  $Q$ . That is,  $w = m\mu_Q$  where  $m$  is constant and  $\mu_Q(T) = \mu(Q \cap T)$  for all measurable  $T$ . Suppose, for the moment, that when  $K$  is large, the  $K$  single facility service regions are all approximately similar (congruous) to a given set  $S$  and have (approximately) equal measure (d-volume)  $\frac{\mu(Q)}{K}$ . This property applies trivially for  $d = 1$  and,

as shown in [HM], [Pa], for  $d = 2$ , but it may be false in general and will not be used in the formal proof. Assume now, without loss of generality, that  $\mu(S) = 1$  and set  $\gamma_d = D_1(\mu_S) = \min_c \int \|x - c\| d\mu_S$ . Our supposition implies that the cost contributed by any single facility service region is approximately  $\gamma_d \frac{m\mu(Q)}{K} \cdot (\frac{\mu(Q)}{K})^{1/d}$ . The terms  $\frac{m\mu(Q)}{K}$  and  $(\frac{\mu(Q)}{K})^{1/d}$  are scaling factors corresponding to the level of the demand within each service region and the size of the region (i.e., how Euclidean distances vary with size scaling). Summing over the  $K$  facilities yields

$$D_K(m\mu_Q) = \gamma_d m\mu(Q) (\frac{\mu(Q)}{K})^{1/d} + o(K^{-1/d}) . \quad (2.1)$$

The term  $o(K^{-1/d})$  tends to 0 faster than  $K^{-1/d}$  as  $K \rightarrow \infty$ .

Next, consider the case where the demand is piecewise uniform, i.e.  $w = \sum_i m_i \mu_{Q_i}$  on a finite collection,  $Q_1, Q_2, \dots, Q_\ell$  of  $d$ -cubes that are slightly removed from each other (i.e.,  $\text{Inf}\{\|x - y\|: x \in Q_i, y \in Q_j\} > 0$  for all  $i, j$ ). Then (see Lemma A7 in Appendix A for a demonstration) for large values of  $K$  no facility will serve demand in more than one of the cubes. Hence for large  $K$ ,  $D_K(\sum_i m_i \mu_{Q_i}) = \sum_i D_{K_i}(m_i \mu_{Q_i})$  for some constants  $K_i$  that denote the number of facilities, out of  $K$ , serving  $Q_i$  for  $i = 1, 2, \dots, \ell$ . A simple argument<sup>3</sup> shows that  $\liminf_{K \rightarrow \infty} \frac{K_i}{K} > 0$  for all  $i = 1, 2, \dots, \ell$  and, consequently,

---

<sup>3</sup>Note that  $D_{K_i}(m_i \mu_{Q_i}) \leq D_K(\sum_j m_j \mu_{Q_j}) \leq \sum_j D_{\lfloor K/\ell \rfloor}(m_j \mu_{Q_j})$  where by (2.1) the left-most expression tends to 0 as fast as  $K_i^{-1/d}$  while the right-most expression tends to 0 as fast as  $K^{-1/d}$  and thus  $K_i$  must grow as fast as  $K$ .

$$\begin{aligned}
 D_K(\sum m_i \mu_{Q_i}) &= \sum_i D_{K_i}(m_i \mu_{Q_i}) = \sum_i (\gamma_d m_i \mu(Q_i) (\frac{\mu(Q_i)}{K_i})^{1/d} + o(K_i^{-1/d})) \\
 &= \gamma_d \sum_i K_i^{-1/d} (m_i^{d/(d+1)} \mu(Q_i))^{(d+1)/d} + o(K^{-1/d}) \quad (2.2)
 \end{aligned}$$

Now consider the following result:

LEMMA 1: [H3, Proposition 13]<sup>4</sup> Suppose  $a_i \geq 0, b_i \geq 0$  for  $i = 1, 2, \dots, \ell$  and  $\alpha < 0$  or  $\alpha > 1$ , then  $\sum_{i=1}^{\ell} a_i^{\alpha} b_i^{1-\alpha} \geq (\sum_{i=1}^{\ell} a_i)^{\alpha} (\sum_{i=1}^{\ell} b_i)^{1-\alpha}$  with equality if and only if  $a_i/b_i$  is constant over  $i$ .

Note:  $a_i/b_i$  is constant for all  $i$  if and only if  $\frac{a_i}{\sum_j a_j} = \frac{b_i}{\sum_j b_j} = \frac{a_i^{\alpha} b_i^{1-\alpha}}{\sum_j a_j^{\alpha} b_j^{1-\alpha}}$

for all  $i = 1, 2, \dots, \ell$ .

Setting  $a_i \equiv K_i, b_i = m_i^{d/(d+1)} \mu(Q_i)$  and  $\alpha = -1/d$  in lemma 1, we have

$$\gamma_d \sum_i K_i^{-1/d} (m_i^{d/(d+1)} \mu(Q_i))^{(d+1)/d} \geq \gamma_d (\sum_i K_i)^{-1/d} (\sum_i m_i^{d/(d+1)} \mu(Q_i))^{(d+1)/d} \quad (2.3a)$$

with equality if and only if

$$\frac{K_i}{K} = \frac{m_i^{d/(d+1)} \mu(Q_i)}{\sum_j m_j^{d/(d+1)} \mu(Q_j)} = \frac{K_i^{-1/d} m_i \mu(Q_i)^{(d+1)/d}}{\sum_j K_j^{-1/d} m_j \mu(Q_j)^{(d+1)/d}} \quad (2.3b)$$

---

<sup>4</sup>Note that if  $p \equiv 1/\alpha, q \equiv 1/(\alpha-1)$ , then  $1/p + 1/q = 1$ . Setting  $f_i \equiv a_i^{\alpha}, g_i b_i^{1-\alpha}$ , we have for  $\alpha < 0$  or  $\alpha > 1$  (i.e.,  $p < 1$  or  $q < 1$ ) that  $\sum_i f_i g_i \geq (\sum_i f_i^p)^{1/p} (\sum_i g_i^q)^{1/q}$  which is an inversion of the usual Holder's inequality.

Since this ratio is asymptotically attainable (take  $K_i = \left\lfloor \frac{m_i^{d/(d+1)} \mu(Q_i)}{\sum_j m_j^{d/(d+1)} \mu(Q_j)} K \right\rfloor$  where  $\lfloor x \rfloor$  is the greatest integer less than or equal to  $x$ , we conclude from (2.2) and (2.3a) that

$$D_K(\sum_i m_i \mu_{Q_i}) = \gamma_d K^{-1/d} (\sum_i m_i^{d/(d+1)} \mu(Q_i))^{(d+1)/d} + o(K^{-1/d}). \quad (2.4)$$

which is a special case of the desired result (2.1).

### 3. MAIN CONVERGENCE RESULTS

Before proceeding to the formal statement of a generalized version of (2.4) and (2.3b), we introduce the following conventions:

-- For any Borel set  $T \subset \mathbb{R}^d$ , let  $w_T$  denote the restriction of  $w$  to  $T$  (i.e.,  $w_T(R) = w(T \cap R)$ )

-- An integer sequence  $\{k_K\}_{K \geq 2}$  is (asymptotically) a T-allocation if  $1 \leq k_K < K$  for all  $K$ , and  $\lim_{K \rightarrow \infty} \frac{D_{k_K}(w_T) + D_{K-k_K}(w_{\mathbb{R}^d - T})}{D_K(w)} = 1$ .

-- Let  $m$  denote the density of the absolutely continuous part of the demand  $w$  (i.e.  $m$  is the Radon-Nikodym derivative  $dw^a/d\mu$  of the absolutely continuous part  $w^a$  of  $w$ ). Although  $m$  depends on the particular  $w$ , in order to keep the notation simple, we do not introduce an explicit index  $w$  for  $m$ .

We are ready now to state our main result:

THEOREM 1: There is a constant  $\gamma_d > 0$ , so that for all  $w$ ,

$$\lim_{K \rightarrow \infty} K^{1/d} D_K(w) = \gamma_d (\int m^{d/(d+1)} d\mu)^{(d+1)/d} \quad (3.1a)$$

Moreover, if  $T$  is any Borel set satisfying  $0 < \int_T m d\mu < \int m d\mu$ , then  $\{k_K\}_{K \geq 2}$  is a  $T$ -allocation if and only if,

$$\lim_{K \rightarrow \infty} \frac{k_K}{K} = \lim_{K \rightarrow \infty} \frac{D_{k_K}^{(w_T)}}{D_K(w)} = \frac{\int_T m^{d/(d+1)} d\mu}{\int m^{d/(d+1)} d\mu} \quad (3.1b)$$

(That is, the densities both of centers and of costs are proportional to  $m^{d/(d+1)}$ .)

To close the gap between the preliminary arguments and a formal proof of Theorem 1, we will (i) prove (3.1a) for demand that is uniform on a cube (i.e., prove (2.1)), without using the assumption that all the single facility regions have some fixed shape asymptotically, and (ii) justify the extension of (2.4) and (2.3b) with  $w = \sum_i m_i \mu_{Q_i}$  to general distributions as in (3.1a, b).

To illustrate the ideas underlying the completion of task (i), consider Figure 1. The square on the left represents the optimal configuration of  $K_0 = 2$  centers in some square with a uniform demand, while the figure to the right represents the partition of that square into 9 scaled down replications of itself (the demand density is unchanged) each with an optimal 2-center configuration.

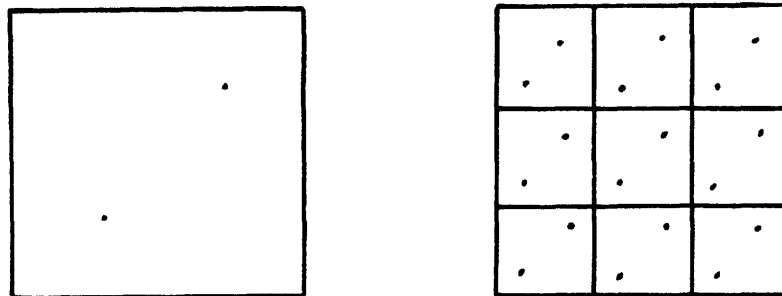


FIGURE 1: Increasing the number of centers 9-fold decreases the cost at least 3-fold.

The cost for each scaled down replication is  $1/27$  of the original cost due to the 3-fold decrease in scale (linear distances) and the 9-fold decrease in area (demand). The cost for a uniform demand on a square with  $2 \cdot 9 = 18$  centers is, then, at most  $(1/27) \cdot 9 = 1/3$  of the cost for such demand with 2 centers. In general, the same argument shows that a  $p^d$ -fold increase in the number of centers,  $K$ , induces at least a  $p$ -fold decrease of the value function  $D_K$  (in our example  $d = 2$ ,  $p = 3$ ). Consequently, the value function is monotonically decreasing and thus convergent on subsequences of  $\{K^{1/d} D_K\}$  of the form  $\{K_q^{1/d} D_{K_q}\}$  where  $K_q = K_0 p^{dq}$  for  $q = 0, 1, 2, \dots$ . The extension, however, of this convergence property to the complete sequence is much more technical and will be pursued only in the proof of Lemma 3 below.

Task (ii) above is carried out in the proof of Lemma 4 which uses some continuity properties of both sides of equation (3.1a). The derivation of these properties is rather technical and deferred to Appendix A.

Reviewing these preliminary arguments suggests that the results of Theorem 1 might be independent of many of the particular features of the Euclidean  $K$ -median problem, and depend instead on rather general properties, like the self similarity of the cube (i.e., a cube is a finite disjoint union of scaled down copies of itself) and the subadditivity, scaling and translation invariance of  $D_K$  (see (P1)-(P4) in Lemma 2 below). These properties are shared by a wide variety of geometrical location problems (e.g., problems with non-Euclidean distances) or even problems involving spatial distribution (allocation) of resources other than static facilities (e.g., salespeople, maintenance/repair personnel, communication (transportation) networking resources like the length of lines (roads)).

In order to facilitate the extension of the results to this broader setting, we abstract from the K-median problem the structure that is essential in establishing Theorem 1.

Lemma 2 below lists those properties of the K-median problem that are relevant to our generalization. Lemmas 3 and 4 apply to any sequence  $\{D_K\}_{K \geq 1}$ , of non-negative functionals on the set of finite positive Borel measures with bounded support, which satisfies some or all of these properties.

Before stating Lemma 2, let us introduce (or reinforce) some notational conventions:

Notation:

- $|w| = w(\mathbb{R}^d)$  denotes the total mass of the demand  $w$ , i.e.,  $|w| = \int dw$ .
- $S(w)$  denotes the support of  $w$ , that is the intersection of all closed sets  $F$  so that  $w(F) = |w|$ .
- $|\mu|_w \equiv \mu(S(w))$ .
- $d_w$  denotes the diameter of  $S(w)$ ; the diameter of a set  $R$  is  $\sup \{ \|x - y\| : x \in R, y \in R \}$ .
- $\lambda^o w$  ( $\lambda > 0$ ) denotes the measure defined  $(\lambda^o w)(R) \equiv w(\frac{1}{\lambda}R)$  where  $\frac{1}{\lambda}R \equiv \{ \frac{1}{\lambda}x : x \in R \}$  for all  $R$ .
- $\lambda w$  is defined, as usual, by  $(\lambda w)(R) \equiv \lambda w(R)$  for all  $R$ .
- $w_{+y}$ , ( $y \in \mathbb{R}^d$ ) is defined for all (Borel)  $R$ , by  $w_{+y} \equiv w(R - y)$  where  $R - y \equiv \{ x - y : x \in R \}$ .
- $w_1 + w_2$  is defined by  $(w_1 + w_2)(R) \equiv w_1(R) + w_2(R)$  for all  $R$ .
- $w_1 \leq w_2$  if  $w_2 - w_1$  is a positive measure.

Definitions

- $R_1$  and  $R_2$  are separated if  $\text{Inf} \{ \|x - y\| : x \in R_1, y \in R_2 \} > 0$ .

-- A d-cell is a cartesian product of  $d$  (one-dimensional) intervals in  $R^d$ .

LEMMA 2:  $D_K(w)$  satisfies the following properties:

- (P1)  $D_{K_1+K_2}(w_1 + w_2) \leq D_{K_1}(w_1) + D_{K_2}(w_2)$  (subadditivity)
- (P2)  $D_K(\lambda^0 w) = \lambda D_K(w)$  (scale linearity)
- (P3)  $D_K(\lambda w) = \lambda D_K(w)$  (demand linearity)
- (P4)  $D_K(w_{+y}) = D_K(w)$  (translation invariance)
- (P5)  $D_{K_1}(w) \leq D_{K_2}(w)$  if  $K_1 \geq K_2$  (monotonicity in resource)
- (P6)  $\text{Sup}\{D_1(w) : |w| \leq 1, d_w \leq 1\} < \infty$  (boundedness)
- (P7)  $D_K(w_1) \leq D_K(w_2)$  if  $w_1 \leq w_2$  (monotonicity in demand)
- (P8) If  $R_1$  and  $R_2$  are bounded and separated,<sup>5</sup> then there is an integer  $M$  with the property that whenever  $S(w_1) \subset R_1$ ,  $S(w_2) \subset R_2$  and  $K \geq 1$ , there are integers  $K_1$  and  $K_2$  such that  $K_1 + K_2 \leq K + M$  and  $D_{K_1}(w_1) + D_{K_2}(w_2) \leq D_K(w_1 + w_2)$   
(induced separability)
- (P9) For the unit d-cube  $S$ ,  $D_K(\mu_S) > 0$  for all  $K$   
(non-triviality)

---

<sup>5</sup>a. It is enough for all later purposes to have (P8) just for separated d-cells (or cubes).

b. While (P8) is stated in terms of a pair of separated sets, it clearly implies by repeated application that for any finite collection of mutually separated sets  $\{R_i\}$  there is a constant  $M$  such that if  $S(w_i) \subset R_i$  for all  $i$  and  $K \geq 1$ , then there are integers  $K_i$  so that  $\sum K_i \leq K + M$  and  $\sum D_{K_i}(w_i) \leq D_K(\sum w_i)$ .



LEMMA 3: For any sequence  $\{D_K\}_{K \geq 1}$  that satisfies (P1)-(P5), there is a constant  $\gamma_d$  (possibly  $-\infty$ ) so that  $\lim_{K \rightarrow \infty} K^{1/d} D_K(m\mu_R) = \gamma_d^m \mu(R)^{(d+1)/d}$  for any d-cube R and constant  $m > 0$ .

Extending this result to general demand distributions  $w$  requires the additional properties (P6)-(P8), of which (P8) seems quite strong (and thus potentially restrictive in other applications). It is instructive, therefore, to show that even without (P8), Lemma 3 extends to demand distributions that are uniform on sets other than cubes. (This extension is not necessary, however, for the derivation of Lemma 4 which involves general demand distributions.)

LEMMA 3\*: (i) Lemma 3 extends without additional assumptions to any region R that is a finite union of d-cells with rational (or commensurate) endpoints. (ii) If (P6) applies as well, then the validity extends to any bounded Lebesgue measurable region R with a boundary of zero Lebesgue measure. (iii) If, in addition, (P7) applies, then the validity extends to any bounded Lebesgue measurable region R.

LEMMA 4: For any sequence of functions  $\{D_K\}_{K \geq 1}$  that satisfies (P1)-(P9), there is a constant  $\gamma_d > 0$  so that (3.1a) holds for all  $w$ . Furthermore, if T is a Borel set with  $0 < \int_T m d\mu < \int m d\mu$ , then  $\{k_K\}_{K \geq 2}$  is a T-allocation if and only if (3.1b) holds.

(Remark: If  $\int_T m d\mu = 0$  or  $\int m d\mu$ , then (3.1b) is necessary, but not sufficient, for a T-allocation.)

Proof of LEMMA 2: (P1)-(P7), (P9) are easy to establish from the definition of  $D_K(w)$ . To see (P8), note that one can always choose  $M$  as the number of balls of diameter  $\frac{1}{2} \inf\{\|x_1 - x_2\| : x_1 \in R_1, x_2 \in R_2\}$  needed to cover

$R_1 \cup R_2$ . Adding any one point of each ball to the optimal center (K-median) set associated with  $D_K(w_1 + w_2)$  does not increase the cost and induces a partition of the (at most  $K + M$ ) centers serving demand in  $R_1 \cup R_2$  into two disjoint sets: those, say  $K_1$  in number, serving demand in  $R_1$  and those, say  $K_2$ , serving demand in  $R_2$ . Consequently,  $D_{K_1}(w_1) + D_{K_2}(w_2) \leq D_K(w_1 + w_2)$ . (Note also that in most cases one can use a constant  $M$  that is significantly lower than the one obtained from the ball cover.) □

Proof of LEMMA 3: Let  $w = \mu_S$  where  $S$  is the unit  $d$ -cube  $[0,1)^d$ . We first prove the convergence of a subsequence of  $\{K^{1/d} D_K(\mu_S)\}_{K \geq 1}$ . As outlined in the preliminary discussion, we use the idea of self-similarity together with the subadditivity, scaling and translation invariance properties.

By properties (P1) and (P4):

$$D_{2^d K}(\mu_S) \leq 2^d D_K(\mu_{\frac{1}{2}S})$$

Noting that  $\mu_{\frac{1}{2}S} = \frac{1}{2^d}(\frac{1}{2} \circ \mu_S)$  we have by (P2) and (P3):

$$D_K(\mu_{\frac{1}{2}S}) = \frac{1}{2^d} \cdot \frac{1}{2} D_K(\mu_S).$$

Consequently,

$$D_{2^d K}(\mu_S) \leq 2^d \frac{1}{2^d} \cdot \frac{1}{2} D_K(\mu_S) = \frac{1}{2} D_K(\mu_S)$$

and

$$(2^d K)^{1/d} D_{2^d K}(\mu_S) \leq K^{1/d} D_K(\mu_S)$$

Hence the subsequence  $\{K^{1/d} D_K(\mu_S)\}_{K=1, 2^d, 2^{2d}, 2^{3d}, \dots}$  is monotonically decreasing and therefore convergent to some constant  $\gamma_d$  (possibly  $-\infty$ ).

We now extend the convergence to the rest of the sequence.

Let  $\ell$  be any positive integer. For every  $K \geq 2^{d\ell}$  there are integers  $q \geq 0$  and  $p$ ,  $2^\ell \leq p < 2^{\ell+1}$  satisfying  $2^{q+\ell} \leq p2^q \leq K^{1/d} < (p+1)2^q \leq 2^{q+\ell+1}$ . Consider now the partitioning of  $S$  into  $p^d$  identical subcubes.

We have by (P1)-(P5)

$$D_K(\mu_S) \leq D_{p^d 2^{dq}}(\mu_S) \leq p^d D_{2^{dq}}(\mu_{(1/p)S}) = D_{2^{dq}}(\mu_S)$$

and thus

$$K^{1/d} D_K(\mu_S) \leq \frac{K^{1/d}}{p} D_{2^{dq}}(\mu_S) < \frac{p+1}{p} 2^{dq} D_{2^{dq}}(\mu_S) \leq (1 + 2^{-\ell}) 2^{dq} D_{2^{dq}}(\mu_S).$$

Since  $2^{dq} D_{2^{dq}}(\mu_S) \rightarrow \gamma_d$  as  $k$  (and therefore  $q$ ) approaches  $\infty$ , we have

$\limsup_{K \rightarrow \infty} K^{1/d} D_K(\mu_S) \leq (1 + 2^{-\ell}) \gamma_d$ . Since  $\ell$  can be arbitrarily large, we have

$$\limsup_{K \rightarrow \infty} K^{1/d} D_K(\mu_S) \leq \gamma_d. \tag{L3.1}$$

On the other hand, suppose that we imbed  $S$  in the larger cube

$(2^{\ell+1})/(p+1)S$ , as shown (for  $d = 2$ ) in Figure 2.

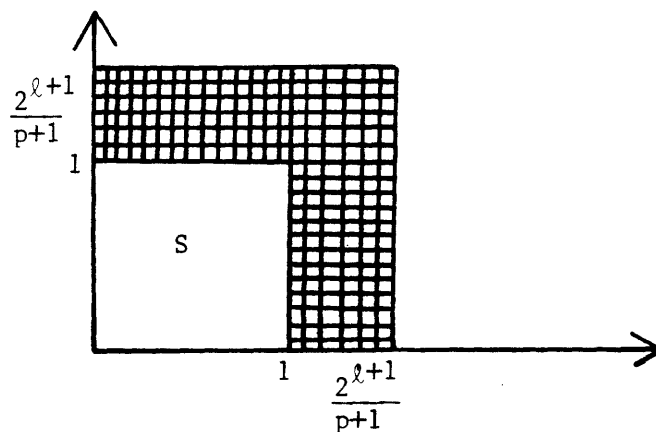


FIGURE 2: Imbedding  $S$  in a larger square (cube).

The set difference between the cubes can be partitioned into  $2^{d(\ell+1)} - (p+1)^d$  cubes, all congruent to  $1/(p+1)S$ . Again by properties (P1) and (P4),

$$D_{2^{d(q+\ell+1)}}(\mu_{(2^{\ell+1}/(p+1))S}) \leq D_{(p+1)^{d_2^{dq}}}(\mu_S) + (2^{d(\ell+1)}) \\ - (p+1)^d D_{2^{dq}}(\mu_{1/(p+1)S})$$

and invoking (P5) and (P2), (P3) we obtain:

$$D_K(\mu_S) \geq D_{(p+1)^{d_2^{dq}}}(\mu_S) \geq \frac{(2^{\ell+1})^{d+1}}{(p+1)^{d+1}} D_{2^{d(q+\ell+1)}}(\mu_S) \\ - \frac{2^{d(\ell+1)} - (p+1)^d}{(p+1)^{d+1}} D_{2^{dq}}(\mu_S).$$

Recalling that  $K^{1/d} \geq p2^q$ , we conclude

$$K^{1/d} D_K(\mu_S) \geq \frac{2^{d(\ell+1)} p}{(p+1)^{d+1}} (2^{q+\ell+1} D_{2^{d(q+\ell+1)}}(\mu_S)) \\ - \frac{(2^{d(\ell+1)} - (p+1)^d) p}{(p+1)^{d+1}} (2^q D_{2^{dq}}(\mu_S)) .$$

But since both  $2^{q+\ell+1} D_{2^{d(q+\ell+1)}}(\mu_S)$  and  $2^q D_{2^{dq}}(\mu_S)$  converge to  $\gamma_d$  as  $K$  (and  $q$ ) approach  $\infty$ , we have

$$\liminf_{K \rightarrow \infty} K^{1/d} D_K(\mu_S) \geq \frac{2^{d(\ell+1)} p}{(p+1)^{d+1}} \gamma_d - \frac{(2^{d(\ell+1)} - (p+1)^d) p}{(p+1)^{d+1}} \gamma_d \\ = \frac{p}{p+1} \gamma_d \geq (1 - 2^{-\ell}) \gamma_d$$

And, since  $\ell$  may be arbitrarily large,

$$\liminf_{K \rightarrow \infty} K^{1/d} D_K(\mu_S) \geq \gamma_d. \tag{L3.2}$$

Combining (L3.1) and (L3.2) gives

$$\lim_{K \rightarrow \infty} K^{1/d} D_K(\mu_S) \geq \gamma_d.$$

Using (P2)-(P4), we conclude for any cube  $R$  and constant  $m \geq 0$

$$\lim_{K \rightarrow \infty} K^{1/d} D_K(m\mu_R) = \gamma_d m^{\mu(R)(d+1)/d} \tag{L3.3}$$

which is equivalent to (3.1a) with  $w = m\mu_R$ . □

Proof of LEMMA 3\*:

(i) We note first that every set  $R$ , that is a finite union of  $d$ -cells with rational endpoints can on one hand be partitioned into

finitely many, say  $\ell$ , congruent cubes, and, on the other hand, can be imbedded in a large cube,  $T$ , with rational endpoints so that the set difference between cube  $T$  and  $R$  can also be partitioned into finitely many, say  $\ell$ , congruent cubes. Properties (P1)-(P5) imply that

$$D_K(m\mu_R) \leq \ell D_{[K/\ell]}(m\mu_{\alpha S}), \quad \text{where } \alpha = \left(\frac{\mu(R)}{\ell}\right)^{1/d}$$

and

$$D_{[K\mu(T)/\mu(R)]}(m\mu_T) \leq D_K(m\mu_R) + q D_{[(1/q)K(\mu(T)/\mu(R)-1)]}(m\mu_{\beta S})$$

where

$$\beta = \left(\frac{\mu(T) - \mu(R)}{q}\right)^{1/d}.$$

Using these two inequalities and the proven convergence for cubic sets one may deduce that (L3.3) is valid in the present case as well.

(ii) Let  $R$  be a bounded Lebesgue measurable set whose boundary has zero Lebesgue measure. As shown by Lemma A6 in Appendix A, for any  $\varepsilon > 0$ , there are sets,  $P$  and  $Q$ , both finite unions of "rational"  $d$ -cells, so that  $P \subset R \subset Q$  and  $\mu(Q - P) < \varepsilon$ . Let  $K_1 = [K\varepsilon]$ ,  $K_2 = K - K_1$  and  $K_3 = K + K_1$  then, by (P1)

$$D_{K_3}(m\mu_Q) - D_{K_1}(m\mu_{Q-R}) \leq D_K(m\mu_R) \leq D_{K_2}(m\mu_P) + D_{K_1}(m\mu_{R-P})$$

By Lemma A1 of Appendix A (which is based on (P1)-(P6)), we know that the upper limits of both  $K_1^{1/d} D_{K_1}(m\mu_{R-P})$  and  $K_2^{1/d} D_{K_1}(m\mu_{Q-R})$  as  $K$  (and thus also  $K_1$ ) tends to  $\infty$ , are bounded by  $\sigma(m\varepsilon) \cdot \varepsilon^{1/d}$  for some  $\sigma$ . Noting that  $K_1/K \rightarrow \varepsilon$ , and using the already proven convergence of  $K_2^{1/d} D_{K_2}(\mu_P)$  and  $K_3^{1/d} D_{K_3}(\mu_Q)$ , we have

$$\begin{aligned} \left(\frac{1}{1+\varepsilon}\right)^{1/d} \gamma_d m\mu(Q)^{(d+1)/d} - \sigma m\varepsilon &\leq \liminf_{K \rightarrow \infty} K^{1/d} D_K(m\mu_R) \\ &\leq \limsup_{K \rightarrow \infty} K^{1/d} D_K(m\mu_R) \leq \left(\frac{1}{1-\varepsilon}\right)^{1/d} \gamma_d m\mu(P)^{(d+1)/d} + \sigma m\varepsilon \end{aligned}$$

and letting  $\varepsilon \rightarrow 0$  establishes convergence for  $K^{1/d} D_K(\mu_R)$ .

(iii) Let  $R$  be any bounded Lebesgue measurable set. For any  $\varepsilon > 0$ , there is a set  $P$  which is a finite union of "rational"  $d$ -cells, so that the symmetric set difference  $P\Delta R \equiv (P - R) \cup (R - P)$  satisfies  $\mu(P\Delta R) < \varepsilon$ ,<sup>6</sup> and thus if  $m_1(x) = m$  for  $x \in R$  and  $m_1(x) = 0$  for  $x \notin R$  and if  $m_2(x) = m$  for  $x \in P$  and  $m_2(x) = 0$  for  $x \notin P$ , then  $\int |m_1(x) - m_2(x)| d\mu = m\mu(P\Delta R) < \varepsilon$ . Now in Lemma A2 of Appendix A we prove, using (P1)-(P7), the  $L^1$  continuity in  $m$  of the lower and upper limits of  $K^{1/d} D_K(w)$ . Thus using the proven convergences of  $K^{1/d} D_K(m\mu_P)$  we may conclude the validity of (L3.3) in the general case.  $\square$

Proof of LEMMA 4:

Consider first the case  $w = \sum_j m_j \mu_{Q_j}$  where  $\{Q_j\}$  is a finite collection of separated cubes: By (P1) and (P8)<sup>7</sup> we have an  $M$  such that for all  $\{K_j\}$  satisfying  $\sum_j K_j = K$ ,  $K_j \geq 1$  and for some  $\{\bar{K}_j\}$  satisfying  $\sum_j \bar{K}_j = \bar{K} \equiv K + M$ ,  $\bar{K}_j \geq 1$

$$(K/\bar{K})^{1/d} \bar{K}^{1/d} \sum_j D_{\bar{K}_j}(m_j \mu_{Q_j}) \leq K^{1/d} D_K(\sum_j m_j \mu_{Q_j}) \leq K^{1/d} \sum_j D_{K_j}(m_j \mu_{Q_j}) \quad (L4.1)$$

---

<sup>6</sup>This follows from the fact that the Boolean ring of finite unions of "rational"  $d$ -cells, generates the Borel  $\sigma$ -algebra (of which Lebesgue  $\sigma$ -algebra is a  $\mu$ -completion and thus can be used to approximate it. ([H1], p. 56.)

<sup>7</sup>See footnote 5.

Now by Lemma 3  $K_j^{1/d} D_{K_j}(m_j \mu_{Q_j}) \rightarrow \gamma_d m_j (\mu_{Q_j})^{(d+1)/d}$  as  $K_j \rightarrow \infty$  and by Lemma 1 (following the preliminary discussion preceding the statement of Theorem 1) we know the minimum of the upper as well as lower bounds in (L4.1) converge to  $\gamma_d (\sum_j m_j^{d/(d+1)} \mu(Q_j))^{(d+1)/d}$  as  $K \rightarrow \infty$ . (Note:  $\bar{K}/K \rightarrow 1$  as  $K \rightarrow \infty$ .) And thus this must be the limit of  $K^{1/d} D_K(\sum_j m_j \mu_{Q_j})$  itself. Finally that  $\gamma_d > 0$  is a straightforward corollary of (P7), (P8) and (P9). □

To extend the result to arbitrary finite Borel measures with bounded support, consider the following lemmas (proven in Appendix A) which facilitate the approximation of any such measure  $w$  by measures of the type  $(\sum_j m_j \mu_{Q_j})$  treated above.

LEMMA A2: Let  $\{D_K\}_{K \geq 1}$  be any sequence that satisfies (P1)-(P7). Then (i) The lower and upper limits of  $K^{1/d} D_K(w)$  do not depend on the singular part of  $w$ . (ii) Let  $Q$  be any bounded set containing the support of  $m$ . Then the lower and upper limits of  $K^{1/d} D_K(w)$  are continuous in  $m$  over  $L^1(Q)$  (the space of integrable functions with support in  $Q$ ).

LEMMA A3: For any bounded and measurable set  $Q \subset \mathbb{R}^d$  and any  $0 \leq q \leq 1$  (in particular  $q = d/(d+1)$ )  $\int m^q d\mu$  is continuous in  $m$  over  $L^1(Q)$ .

LEMMA A4: The set of functions of the form  $\sum_{i=1}^{\ell} m_i l_{Q_i}$ , where  $l_{Q_1}, l_{Q_2}, \dots, l_{Q_\ell}$  are indicator functions of separated cubes, is dense in  $L^1(\mathbb{R}^d)$ .

These lemmas imply that the validity of (3.1a) extends to arbitrary  $w$ .



Now to prove the second part of the assertion of Lemma 4, assume first that  $\lim_{K \rightarrow \infty} \frac{k_K}{K} = \frac{A_T}{A}$  where  $A_T \equiv \int_T m^{d/(d+1)} d\mu$  and  $A = \int m^{d/(d+1)} d\mu$ . By (3.1a) we have, as  $K \rightarrow \infty$ :

$$\frac{D_{k_K}(w_T) + D_{K-k_K}(w_{R^{d-T}})}{D_K(w)} = \frac{\left(\frac{k_K}{K}\right)^{-1/d} k_K^{1/d} D_{k_K}(w_T) + \left(1 - \frac{k_K}{K}\right)^{-1/d} D_{K-k_K}(w_{R^{d-T}})}{K^{1/d} D_K(w)}$$

$$\rightarrow \frac{\left(\frac{A_T}{A}\right)^{-1/d} \gamma_d^{A_T^{(d+1)/d}} + \left(\frac{A - A_T}{A}\right)^{-1/d} \gamma_d^{(A - A_T)^{(d+1)/d}}}{\gamma_d^{A^{(d+1)/d}}} = 1.$$

That is,  $\{k_K\}_{K \geq 2}$  is a T-allocation.

Conversely, assume that  $\{k_K\}_{K \geq 2}$  is a T-allocation.

We first prove:

Claim:  $0 < \liminf_K \frac{k_K}{K} \leq \limsup_K \frac{k_K}{K} < 1$

Proof:

(i) If  $\liminf k_K = M < \infty$ , then there is a subsequence of  $\{K^{1/d} D_{k_K}(w_T)\}_{K \geq 2}$  that converges (diverges) to 0,  $+\infty$  or  $-\infty$  according to  $D_M(w_T)$  being 0, positive or negative.

(ii) If  $\lim_K k_K = \infty$  but  $\lim_K \frac{k_K}{K} = 0$  then by (3.1a)  $\{K^{1/d} D_{k_K}(w_T)\}_{K \geq 2}$  diverges to  $+\infty$ .

(iii) By (3.1a)  $\lim_K K^{1/d} D_K(w) = \gamma_d^{A^{(d+1)/d}}$  and if  $\lim_K \frac{k_K}{K} = 0$ , then

$$\lim_K K^{1/d} D_{K-k_K}(w_{R^{d-T}}) = \lim_K \left(\frac{K}{K - k_K}\right)^{1/d} (K - k_K)^{1/d} D_{K-k_K}(w_{R^{d-T}})$$

$$= \gamma_d^{(A - A_T)^{(d+1)/d}}$$

Thus if  $k_K/K \rightarrow 0$ , then by (i)-(iii)

$$\frac{K^{1/d} D_{k_K}(w_T) + K^{1/d} D_{K-k_K}(w_{R^d-T})}{K^{1/d} D_K(w)} \rightarrow \frac{z + \gamma_d (A - A_T)^{(d+1)/d}}{\gamma_d A^{(d+1)/d}}$$

where  $z$  is  $0$ ,  $+\infty$ , or  $\infty$ , making the right hand side of the last expression, everything but  $1$ , which contradicts the assumption that  $\{k_K\}$  is a

$T$ -allocation. We deduce then, that  $\liminf_K \frac{k_K}{K} > 0$ .

Since we could reverse the roles of  $T$  and  $R^d - T$ , we have (recalling that  $A_T < A$ ) also  $\limsup_K \frac{k_K}{K} < 1$  (QED claim). It follows, then, using (3.1a), that

$$K^{1/d} D_{k_K}(w_T) = \left(\frac{k_K}{K}\right)^{-1/d} \gamma_d A_T^{(d+1)/d} + o(1)$$

$$K^{1/d} D_{K-k_K}(w_{R^d-T}) = \left(1 - \frac{k_K}{K}\right)^{-1/d} \gamma_d (A - A_T)^{(d+1)/d} + o(1).$$

And thus

$$\frac{D_{k_K}(w_T) + D_{K-k_K}(w_{R^d-T})}{D_K(w)} = \frac{\left(\frac{k_K}{K}\right)^{-1/d} A_T^{(d+1)/d} + \left(1 - \frac{k_K}{K}\right)^{-1/d} (A - A_T)^{(d+1)/d} + o(1)}{A^{(d+1)/d} + o(1)}$$

and by Lemma 1 (and the continuity of the numerator of the right hand side in  $k_K/K$ ), the last expression is equal to  $1 + o(1)$  if and only if

$\frac{k_K}{K} = \frac{A_T}{A} + o(1)$ . We skip the straightforward proof that  $\frac{k_K}{K} \rightarrow \frac{A_T}{A}$  is

equivalent to  $\frac{D_{k_K}(w_T)}{D_K(w)} \rightarrow \frac{A_T}{A}$ . □

#### 4. RESTRICTED MEDIANS

Consider a variation of the K-median problem in which the set of facility locations is restricted to (the closure of) the set of customer locations, i.e.,  $C \subset S(w)$ . We define:

$$\bar{D}_K(w) = \text{Min} \{D(w,C) : C \subset S(w), |C| \leq K\}$$

The restricted median problem reflects the fact that often a facility must be located in or next to an urban or other inhabited area. It is also the direct analog of the graph theoretic K-median problem in which the centers (or medians) are a subset of the nodes (customer locations).

Observe that with the exception of the monotonicity (in demand) property, (P7),  $\bar{D}_K(w)$  satisfies all of the properties listed in Lemma 2. Lemma 3\* implies convergence of  $K^{1/d} \bar{D}_K(\mu_R)$  when R has a boundary of zero Lebesgue measure. And since for a square R  $\bar{D}_K(\mu_R) = D_K(\mu_R)$ , the convergence is to the same limit. In fact, the results of Theorem 1, in their full generality, apply to  $\bar{D}_K(w)$  with the same constant  $\gamma_d$  used for  $D_K(w)$ . The technicalities of proof are deferred to Appendix B.

#### 5. EXTENSIONS AND MODIFICATIONS

The K-median problem is but one type of facility location problem. A broad class of models are obtained by permitting costs to be nonlinear in the customer-facility distances or in the demand, by imposing additional costs or constraints (e.g., set-up costs, capacitated facilities), and by permitting non-isotropic (i.e., non-Euclidean) metrics.

The facility location problem is a decision problem requiring the allocation of scarce resources (the facilities) over space. Other spatial allocation problems may be concerned with the distribution of various resources like a bus fleet, police patrol cars, repair and maintenance personnel, salespeople and so forth (that may be viewed as mobile facilities) to serve some spatially distributed demand.

Sometimes the resources in question are continuous, as in the following modification of Steiner's street network (spanning-tree) problem: Find the connected network with total length no more than  $K$  that minimizes the average distance between a demand point (i.e. a customer) and the point on the network that is closest to it. (This problem is relevant, for example, to the design of a public transportation system like a subway). The resource,  $K$ , here is the total length allowed for the network.

### 5.1 Nonlinear (Homogeneous) Costs

In many facility location problems the cost of serving a customer at  $x$  by a facility at  $c$  is not proportional to distance  $\|x - c\|$ . Usually the cost is some monotonically increasing function of  $\|x - c\|$ . The simplest form of nonlinearity akin to our analysis is when the cost is proportional to  $\|x - c\|^\alpha$  with  $\alpha > 0$ . In such cases, the scaling property (P2),  $D_K(\lambda^0 w) = \lambda D_K(w)$  (scale linearity), becomes:

$$(P2_\alpha) \quad D_K(\lambda^0 w) = \lambda^\alpha D_K(w) \quad (\text{scale homogeneity}).$$

Similarly, in some facility location problems the cost of serving two customers at  $x$  by a facility in  $c$  is less (or more) than twice the

cost of serving a single customer. That is, the cost is not necessarily linear in the demand. Consider, for example, the problem of designing a pipeline network to distribute some fluid, say natural gas, from a central facility to customers with certain flow rate requirements. Assume that the fluid is incompressible and that there are no viscosity losses or pumping stations between the facility and the customers so that the fluid velocity is fixed and identical throughout the system. The flow rate in a pipe, then, is proportional to the area of its section, that is to the square of its diameter. Suppose now that the pipes are made of steel and the wall thickness is the same for all pipes. Then the amount of steel per mile pipe length is proportional to the square root of the flow rate in the pipe. If the design objective is to satisfy demand while minimizing the amount of steel needed, then scaling the demands, or flow rates, 2-fold results, for a fixed velocity, in only a  $1.414 = 2^{\frac{1}{2}}$ -fold increase in the amount of steel required, which is the cost here. (It is possible to show that the scaling of the demand does not affect the configuration of the optimal distribution network.)

A similar example arises when wiring a population of customers to a set of  $K$  exchanges (e.g. central processors). The wires are all of the same kind. Each wire lies in a cable that carries one or more wires. The protective coating of the cables has some fixed thickness. Three possible objectives for minimization are (1) the total wire length, which indicates the amount of copper required; (2) the total cable length which indicates, say, the amount of labor in laying the cable; (3) the surface area of the cables which indicates the amount of coating required. Note that in all three cases the cost of a cable link is proportional to  $\ell \cdot f^{\beta}$  where  $\ell$  is the length of the link,  $f$  ("the flow")

is the number of wires in the link, and  $\beta$  is 1, 0 or  $\frac{1}{2}$ , respectively, for the three objectives.

To illustrate the different geometries that result from each of the above scenarios consider Figure 3, where we have one facility serving five identical customers. Note that for  $\beta \geq 1$  (i.e., costs are linear in link length and convex in flow) an optimal Euclidean network will use the straight line segments connecting sources to sinks, while for  $\beta = 0$  (i.e. costs are linear in link length but independent of flow) an optimal Euclidean network will consist of a Steiner (spanning) tree (or forest). Optimal Euclidean networks for  $0 < \beta < 1$ , will have an intermediate nature, with "forking" points in the network other than sources or sinks as in a Steiner tree, but the angle between the two incoming (or outgoing) flows at such a "forking" point joining 3-links will be smaller than  $120^\circ$ . When the two incoming flows are equal, the obtuse angle between each of them and the outgoing flow has a cosine equal to  $1/2^{1-\beta}$ . (See Appendix C for a calculation of forking angles in optimal Euclidean flows.) It is possible to show that scaling the demands (supplies)

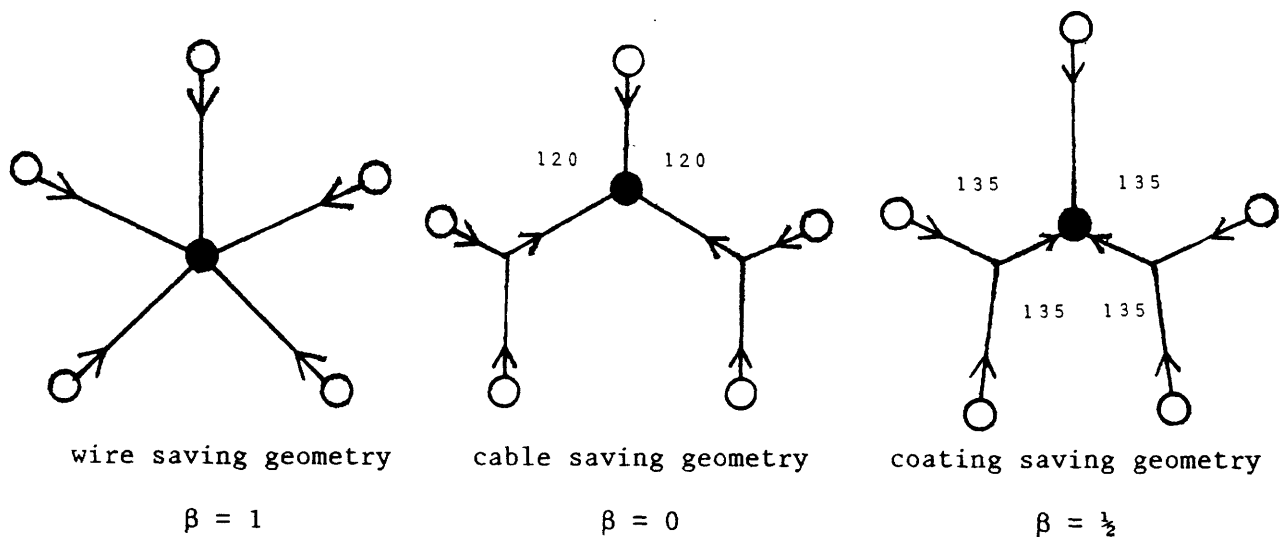


FIGURE 3: Economic geometries for the wiring problem.

$\lambda$ -fold in a network where all the arcflow costs are homogenous in the flow with some degree  $\beta$ , will not change, beyond scaling, the optimal flow, and will scale the cost  $\lambda^\beta$ -fold.

The only difficulty in adapting the analysis of Section 2 to this type of problem is that, in contrast to the K-median problem (i.e.  $\beta = 1$ ), there seems to be no simple intuitive cost definition for continuous, rather than finite discrete, demand distributions. One may, of course, take sequences of successive discrete approximations that converge to a given, say uniform, continuous demand. It is not clear, however, that all such discrete approximations will converge to a unique limit. We shall then, in the present discussion, bypass the fine points of defining a cost for continuous demand distributions. (Note that if we would have constrained customers to use the straight line segment connecting them to the facility we would have no difficulty in defining the problem for (absolutely) continuous demand. Just substitute  $m^\beta d\mu$  instead of  $dw$  in the definitions of  $D(w,C)$ .) There are, then, location problems where the linearity in the demand property (P3),  $D_K(\lambda w) = \lambda D_K(w)$ , becomes:

$$(P3_\beta) \quad D_K(\lambda w) = \lambda^\beta D_K(w) \quad (\text{demand homogeneity})$$

The case  $0 \leq \beta < 1$  will correspond as in the previous example to an economy of scale, while  $\beta > 1$  will correspond to congestion effects. For example, if we have many customers concentrated in some vicinity, who have to travel to the same facility, then busing or car-pooling introduces economies of scale, while everyone using his own car adds congestion as the customers interfere with each other on the presumably capacitated transportation links to the facility.

Using the homogenous scaling rules  $(P2_\alpha)$  and  $(P3_\beta)$  rather than the linear rules (P2) and (P3), the results of Lemma 4 (Theorem 1) become:

For  $\beta + \alpha/d > 1$  and for all  $w$ ,

$$\lim_{K \rightarrow \infty} K^{\beta + \alpha/d - 1} D_K(w) = \gamma \left( \int m^{1/(\beta + \alpha/d)} \right)^{\beta + \alpha/d} \quad (5.1a)$$

$$\lim_{K \rightarrow \infty} \frac{k_K(T)}{K} = \frac{\int_T m^{1/(\beta + \alpha/d)} d\mu}{\int m^{1/(\beta + \alpha/d)} d\mu} \quad (\text{provided } \gamma > 0) \quad (5.1b)$$

The proof of these results follows closely that of Theorem 1. Note that if  $\beta + \alpha/d < 1$ , and we assume that a non-vanishing fraction of the single facility subregions has area of order  $\frac{\mu(R)}{K}$ , then the coefficient of  $K$  in this asymptotic expression for  $D_K(m\mu_R)$  is positive thus  $D_K(m\mu_R)$  increases as  $K \rightarrow \infty$  which is of course impossible, by property (P5). Consequently, most of the facilities will not be used by the optimal solution. Or, in other words, the economy of scale discourages the use of too many facilities.

As we have seen in the wiring example, in some cases the cost will be a linear combination of homogenous costs. Now assume that the asymptotic center configuration for uniform demand distributions is the same for all the homogenous costs. This is indeed the case for  $d = 2$  (see [HM]). Assume, also, that the demand distribution is absolutely continuous with a density  $m$ , and that the centers are distributed according to the (relative) density  $\kappa$  (i.e., the number of centers out of a total of  $K$  that are located in  $T$  is approximately  $K \int_T \kappa d\mu$ ). Assuming that the asymptotic center configuration is at least locally optimal (i.e., is



the same as for uniform demand), we may conclude that if we have several, say  $\ell$  homogenous costs, then (for a given  $\kappa$  and a large value of  $K$ ),

$$\text{total cost} \cong \sum_{i=1}^{\ell} \gamma^{(i)} K^{1-\beta_i} - \frac{\alpha_i}{d} \int m_i \kappa^{\beta_i} K^{1-\beta_i} - \frac{\alpha_i}{d} d\mu \quad (5.2)$$

where the  $\gamma^{(i)}$ s are "appropriate" constants.

It follows that as  $K \rightarrow \infty$ , the cost component with the smallest  $\beta + \alpha/d$  becomes dominant, and if one, accordingly, neglects the other components and minimizes the cost subject to the constraint  $\int \kappa d\mu = 1$  ( $\kappa \geq 0$ ), then (5.2) reduces to (5.1a). If, however,  $K$  is assumed large enough for the validity of the asymptotic approximation (5.2), but not large enough to neglect the terms with order higher than  $\beta + \alpha/d$ , one may still look for a relative center density,  $\kappa$  that minimizes the right hand side of (5.2), although, as opposed to the purely homogenous case, the optimal density  $\kappa$  will depend on  $K$  and its derivation is not so simple.

## 5.2 Facility Set-up (Fixed Charge) Costs

In most facility location problems, one should account for costs that do not depend on the distances between the facilities and their customers (i.e., costs other than what we may call "geometrical" costs), such costs are usually due to the cost of opening and operating a facility at a certain location to serve a certain volume of customers, without regard to distances of these customers to this location. In the wiring example, these are the costs of building (and operating) an exchange of a certain capacity of lines at each given point. Usually such costs will contain a fixed charge component, that is some cost that does not depend

on the size of the switchboard of the exchange, and another component depending on that size (that is, on the capacity of the exchange).

In general, then, we have (assuming that locally the demand is evenly distributed among the facilities) a function  $g(x, \frac{m(x)}{K(x)})$  so that the additional cost is:

$$K \int g(s, \frac{m}{K}) \kappa d\mu \quad (5.3)$$

One may, of course, add this expression to the previously discussed asymptotic approximation to the "geometrical" costs, and minimize over  $K$  and  $\kappa$ , provided of course the resulting  $K$  is large enough<sup>8</sup> to support the asymptotic approximation.

In the simplest case of a fixed charge problem  $g(x, \frac{m}{K}) = F$ . Suppose, in addition, that the geometrical cost is homogenous (with scale factor  $\beta + \alpha/d > 1$ ). The combined cost will be approximately

$$K^{-(\beta+\alpha/d-1)} \gamma (\int m^{1/(\beta+\alpha/d)} d\mu)^{\beta+\alpha/d} + F K$$

and a short calculation yields:

$$k^*(T) \cong \left( \frac{(\beta + \alpha/d - 1)\gamma}{F} \right)^{1/(\beta+(\alpha/d))} \int_T m^{1/(\beta+(\alpha/d))} d\mu$$

where  $k^*(T)$  is the number of centers to be allocated to the subset  $T$ . This expression makes sense, of course, only if the resulting total number of centers,  $K^* = k^*(R^d)$  is large enough to make the approximation valid.

---

<sup>8</sup>And one should add "and the resulting  $\kappa$  is 'smooth' enough."

### 5.3 Capacitated Facilities

Suppose that any single facility is able to handle only some limited number of customers, and/or, on the other hand, suppose that a facility may be established only if some minimal number of customers is guaranteed to use it.

Using the asymptotic approximation, this requirement can be written in form of the constraint  $L \leq \frac{m(x)}{K\kappa(x)} \leq U$  for all  $x$  (we assume an absolutely continuous demand), where  $L$  and  $U$  are some positive constants. Note that as  $K \rightarrow \infty$  the upper capacity constraint becomes nonbinding whereas the lower capacity constraint becomes infeasible unless  $L = 0$ . As before one may still use this approximation, assuming a finite though large value of  $K$ . We shall, however, modify the capacity constraints as follows.

$$\ell \leq \frac{1}{|w|} \frac{m(x)}{\kappa(x)} \leq u \quad \text{for all } x \in S(w) \quad (5.4)$$

with  $0 \leq \ell \leq 1$  and  $u \geq 1$ .

(Recall that  $|w| = \int m d\mu$  since we assumed that  $w$  is absolutely continuous, and note that the average of  $(1/|w|)(m/K)$  over the facilities is  $\int (1/|w|)(m/\kappa)\kappa d\mu = 1$ .)

Assume now, without loss of generality, that  $|w| = \int m d\mu = 1$ . Consider the case of a homogenous cost with scale factor  $\beta + \alpha/d > 1$  and with capacity constraints (5.4). To find the optimal relative center distribution,  $\kappa$ , we have to solve the problem:

$$\begin{aligned} &\text{Minimize} && \int m^\beta \kappa^{1-\beta-(\alpha/d)} d\mu \\ &\text{Subject to:} && \int \kappa d\mu = 1 \\ &&& \frac{1}{u} m(x) \leq \kappa(x) \leq \frac{1}{\ell} m(x) \quad \text{for all } x \in S(w) \end{aligned} \quad (5.5)$$

It may be shown that an optimal center density,  $\kappa^*$  exists. Consider such optimal  $\kappa^*$  and let  $\bar{A} \equiv \{x \in S(w) : \kappa^*(x) = 1/u m(x)\}$ ,

$$\underline{A} \equiv \{x \in S(w) : \kappa^*(x) = \frac{1}{\ell} m(x)\} \quad \text{and} \quad B \equiv \{x \in S(w) : \frac{1}{u} m(x) < \kappa^*(x) < \frac{1}{\ell} m(x)\}$$

It could be argued that  $\kappa^*$  depends on  $x$  through  $m$  alone, and that it should be continuous in  $m$ . (Suppose we take two small neighborhoods with roughly the same, approximately constant, demand density and the same area, then they should approximately have the same number of centers.) It is also clear that the restriction of  $\kappa^*$  to  $B$  should be optimal (up to a constant factor) with respect to the incapacitated problem for the restriction of  $m$  to  $B$ , i.e. there is a constant,  $\beta$ , so that  $\kappa^*(x) = \beta m^\theta(x)$  to all  $x \in B$ , where  $\theta \equiv \beta + \alpha/d$ . Recalling that  $\kappa^*$  should be continuous in  $m$ , we conclude that:

$$\kappa^*(x) = \begin{cases} \frac{1}{\ell} m(x) & \text{if } m(x) < \underline{m} \\ \beta m^\theta(x) & \text{if } \underline{m} \leq m(x) \leq \bar{m} \\ \frac{1}{u} m(x) & \text{if } m(x) > \bar{m} \end{cases} \quad (5.6)$$

where  $\underline{m}$  and  $\bar{m}$  are solutions of  $1/\ell \underline{m} = \beta \underline{m}^\theta$  and  $1/u \bar{m} = \beta \bar{m}^\theta$ . (5.7)

Assume now that  $u < \infty$ . Then  $\bar{m} < \infty$  and by (5.7), we have:

$$\beta = \frac{1}{u} \bar{m}^{-1-\theta} \quad ; \quad \underline{m} = \lambda \bar{m} \quad \text{where } \lambda \equiv (\ell/\mu)^{1-\theta} \quad (5.8)$$

Making these substitutions in (5.7) and integrating, gives:

$$\int \kappa^* d\mu = \frac{1}{\ell} \int_{\{x: m \leq \lambda \bar{m}\}} m d\mu + \frac{1}{u} \int_{\{x: \lambda \bar{m} < m < \bar{m}\}} \bar{m}^{-1-\theta} m^\theta d\mu + \frac{1}{u} \int_{\{x: m \geq \bar{m}\}} m d\mu = 1 \quad (5.9)$$

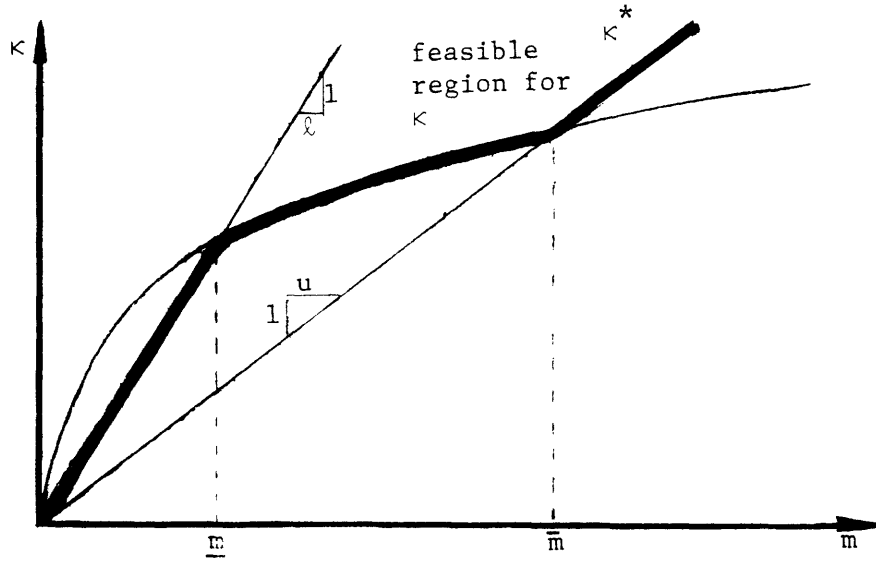


FIGURE 4: Center density as a function of demand density.

which determines the value of  $\bar{m}$  and thus also the values of  $\beta$  and  $\underline{m}$  in (5.6) above. To see that (5.9) has a unique solution for  $\bar{m}$ , note that the left hand side of (5.9) is monotonically increasing and continuous in  $\bar{m}$ . Moreover, for  $\bar{m} = 0$  the left hand side is 0 while for  $\bar{m} \rightarrow \infty$  it converges to  $\int 1/\ell \, m \, d\mu = 1/\ell > 1$ . Therefore by the mean value theorem there exists an  $\bar{m}$  so that (5.9) holds and by monotonicity this  $\bar{m}$  is unique. It is also quite simple to find  $\bar{m}$  (by binary search for example).

If  $u = \infty$ , but  $\ell > 0$  then  $\bar{m} = \infty$ ,  $\underline{m} > 0$  and  $\beta = 1/\ell \, \underline{m}^{1-\theta}$  and (5.9) can be modified to

$$\int \kappa^* d\mu = \frac{1}{\ell} \int_{\{x:m < \bar{m}\}} m d\mu + \frac{1}{\ell} \int_{\{x:m \geq \bar{m}\}} \underline{m}^{1-\theta} m^\theta d\mu = 1 \quad (5.10)$$

#### 5.4 Non-isotropic (Euclidean) Metrics

In Sections 3, 4 and in this section (with the exception of the explicit description of the economic geometries for the wiring example),

we have not used the isotropy of the Euclidean distances and thus the asymptotic results of Sections 3, 4 and their modification in the present section are valid (up to the constant  $\gamma$ ) for other metrics such as the  $\ell^1$  (rectilinear) metric, which may be more appropriate in urban areas.

## 6. FRACTAL MODELS FOR CLUSTERED DEMAND

Finally, we would like to make some comments on the case when the customer population (the demand) is highly clustered, that is  $S(w)$ , the support of the demand is essentially singular. Note that (3.1a) and (5.1a) will still hold, but with a null right hand side, which does not reflect any information beyond the fact that the cost  $D_K$  tends to zero at a rate faster than  $K^{-1/d}$  (or  $K^{1-\beta-a/d}$ ). What, then, is the asymptotic behavior of the cost for such distributions? The question is, of course, senseless in the case where the demand is concentrated on a finite number of points since the cost can be nullified by a finite number of facilities placed at these points. For the case where the support is infinite, although singular, one may suggest "fractal" or fractional dimensional models. In his book "Fractals", [M2], Mandelbrot advocates such models for, among many other things, the distribution of stellar matter in the universe. Though intergalactic facility location does not seem to be an urgent concern, it is quite plausible that population habitation and dispersion processes (and the resulting customer distributions) tend to follow a hierarchical recurrent pattern which a fractal may model quite better than a continuous distribution.

We shall not dwell here on the definition of fractal sets or fractional (Housdorff-Besicovitch) dimension of sets or measures but,

instead, refer the interested reader to references [Ma] and [Fl]. We shall, however, demonstrate by a simple example what kind of modified results can be expected for this type of distributions.

Consider a unit demand, distributed (uniformly) on Cantor's triadic (deleted middle thirds) set  $C$  (defined as the largest subset of  $[0,1]$  satisfying the recursion  $C = \frac{1}{3}(C \cup (C+2))$ , i.e.  $C = \{x \in [0,1] : 3x \in C \text{ or } 3x-2 \in C\}$ ). This demand is given by the distribution  $w$ , with total mass  $|w| = 1$  on the interval  $[0,1]$ , that is recursively defined by

$$w = \frac{1}{2} \left( \frac{1}{3} \circ (w + w_{+2}) \right) \tag{6.1}$$

Recall now that in the proof of Theorem 1 for the case of uniform demand on the  $d$ -cube  $S = [0,1]^d$  (i.e., in the proof of Lemma 3), we used the fact that, for any  $\ell = 2, 3, \dots$

$$\mu_S = \frac{1}{\ell^d} \left( \frac{1}{\ell} \circ \sum_{i=0}^{\ell^d-1} \mu_{S+x_i} \right) \tag{6.2}$$

where the  $x_i$ 's are the integer points in the cube  $[0,\ell]^d$ .

From (6.2) with  $\ell = 2$ , and properties (P1)-(P4) of Lemma 2, we deduced the convergence of the subsequence  $\{K^{1/d} D_K(\mu_S)\}_{K=1, 2^d, 2^{2d}, \dots}$  to some  $\gamma_d \geq 0$ . In a similar fashion, we may deduce from (6.1) and (P1)-(P4) the convergence of  $\{K^{1/\delta} D_K(w)\}_{K=1, 2, 2^2, \dots}$  where  $\delta = \log_3 2$ . (Comparing (6.1) and (6.2) observe that  $d \equiv \log_2 \ell^d$ . Having  $\delta = \log_3 2$ , play the same role that  $d$  played for the uniform distribution on a solid  $d$ -dimensional cube explains, in a sense, the rationale of regarding  $\delta$  as the (fractional) dimension of the Cantor set and the Cantor distribution.)

Unlike the case of the uniform distribution on the solid cube, however, the convergence of  $\{K^{1/\delta} D_K(w)\}_{K=1,2,2^2,\dots}$  does not extend to the whole sequence. One may explain this difference by the fact that the self-similarity property (6.2) is somewhat stronger than the self-similarity property (6.1). Compare for example (6.1) with (6.2) for the case  $d = 1$ , i.e., when  $S$  is the interval  $[0,1)$ . Note that while the interval  $[0,1)$  can be partitioned to any number  $\ell$  of identical scaled down and translated copies of itself, the Cantor set can be partitioned only to any power of 2, of identical scaled down and translated copies of itself.

Nonetheless, using the fact that there is convergence along subsequences of the form  $K = p, 2p, 2^2p, \dots$  for  $p \geq 1$ , we may suggest a modified convergence result of the form  $\frac{K^{1/\delta} D_K(w)}{g(K/(2^{\lfloor \log_2 K \rfloor}))} \rightarrow \gamma$  where  $g$  is some (to be specified) real valued function defined on  $[1,2)$  with  $g(1) = 1$ .

An explicit computation (Appendix D) yields,  $\gamma = 1/3$  and  $g(x) = x^{1/\delta} (\frac{5}{3} - \frac{2}{3} x)$ .

Similar conclusions may be obtained for fractal demand distribution lying in  $\mathbb{R}^d$  for  $d > 1$ , and we conjecture that similar results may be obtained for randomly generated fractal distributions.



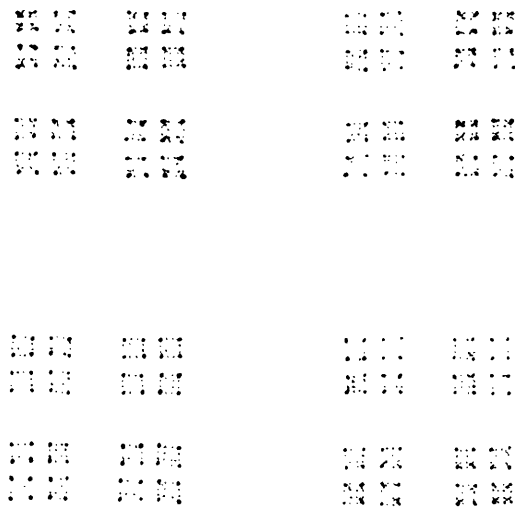


FIGURE 5: A fractal set in the plane (the Cartesian square of the Cantor set)

## 7. SUMMARY

In this paper we have established asymptotic properties of the K-median problem for deterministic and absolutely continuous demand distributions. We demonstrated the convergence of  $K^{1/d} D_K(w)$  for a d-dimensional version of the problem defined by abstracting quite general properties satisfied by the K-median problem. Our discussion has introduced extensions and modifications of this result for different versions of geometrical location problems where costs are not necessarily linear in distances or in the demand, where set-up costs or capacities for facilities may be included in the model, where distances are not necessarily Euclidean, and where the demand distribution (customer population) is highly clustered in an hierarchical manner that admits so-called fractal modeling.

The applications of the results in this paper lie in providing "fast" approximate formulas for the cost and for the allocation of facilities in large scale location problems, especially for preliminary design purposes. They may also be used in heuristics that utilize (locally) the hexagonal partition property ([F2], [HM], [P1]).

In a companion paper [Ha] we provide a rigorous justification to the use of continuous demand distribution, as a "smoothed" representation of basically discrete customer populations. We show that if  $N$  customers are independently located at random according to some probability distribution  $p$  then as  $N/K \rightarrow \infty$  (whether  $K$  is fixed or  $K \rightarrow \infty$  simultaneously), the average cost and the center locations to the associated  $K$ -median problem coincide asymptotically (with probability 1) with those of the  $K$ -median problem with the probability measure  $p$  as the demand.

8. REFERENCES

- [Co] Courant, R. and H. Robbins (1947). What is Mathematics? Oxford University Press.
- [F1] Federer, H. (1969). Geometric Measure Theory. Springer, New York.
- [F2] Fejes Tóth, L. (1959). "Sur la Représentation d'une Population Infinie par un Nombre Fini d'Éléments," Acta Math. Acad. Sci. Hungar, 10, 299-304.
- [F3] \_\_\_\_\_ (1953). Lagerungen in der Ebene, auf der Kugel und im Raum. Berlin-Göttingen-Heidelberg.
- [Ha] Haimovich, M. (1985). "Large Scale Geometric Location Problems II: Stochastic Limit Properties." Working Paper OR 131-85, Operations Research Center, M.I.T., January.
- [HM] \_\_\_\_\_ and T. L. Magnanti (1985). "Extremum Properties of Hexagonal Partitioning and the Uniform Distribution in Euclidean Location." Working Paper OR 132-85, Operations Research Center, M.I.T., January.
- [H1] Halmos, P. R. (1950). Measure Theory. Springer Verlag, New York-Heidelberg-Berlin (1974).
- [H2] Handler, G. Y., and P. B. Mirchandani (1979). Location on Networks, Theory and Algorithms. MIT Press, Cambridge, Massachusetts.
- [H3] Hardy, G. H., Littlewood, J. E., and G. Pólya (1934). Inequalities. Cambridge.
- [K1] Kuenne, R. E., and R. M. Soland (1972). "Exact and Approximate Solutions to the Multi-Source Weber Problem," Math. Progr. 3, 193-209.
- [K2] Kuhn, H. W. (1967). "On a Pair of Dual Nonlinear Problems," in Methods of Nonlinear Programming (ed. J. Abadie). North Holland, Amsterdam.
- \_\_\_\_\_ (1974), "'Steiner's" Problem Revisited," in Studies in Optimization, (eds. R.W. Cottle and G.B. Dantzig).
- [Ma] Mandelbrot, B. B. (1977). Fractals: Form Chance and Dimension. Freeman, San Francisco.
- [Pa] Papadimitriou, C. H. (1981). "Worst Case and Probabilistic Analysis of a Geometric Location Problem." SIAM J. Comput., 10, 542-557.
- [Ru] Rudin, W. (1966). Real and Complex Analysis. McGraw-Hill, New York.

[We] Weber, A. (1909). Über den Standort der Industrien. Tübingen.

APPENDIX A

LEMMA A1: For any  $\{D_K\}_{K \geq 1}$  that satisfies (P1)-(P6), there is a constant  $\sigma \geq 0$  so that:

$$\limsup_{K \rightarrow \infty} K^{1/d} D_K(w) \leq \sigma |w| |\mu|_w^{1/d}$$

Proof: Let  $R = S(w)$ ,  $|\mu|_w = \mu(R)$ ,  $|w| = w(R)$ .

Claim: There is a set  $P$  which is a finite union of disjoint cubes so that:

$$\mu(P \Delta R) < \varepsilon \quad \text{and} \quad w(P \Delta R) < \varepsilon^2 \quad (\text{where } P \Delta R \equiv (P - R) \cup (R - P))$$

Proof of claim:

By the regularity of  $\mu$ , we know that there exist a compact set  $Q_1$  and an open set  $V_1$  so that  $Q_1 \subset R \subset V_1$  and  $\mu(V_1 - Q_1) < \varepsilon$ .

Similarly by the regularity of  $w$  (recall that any finite Borel measure on a Euclidean space is regular [Ru], Th. 2.18) there exists a compact set  $Q_2$  and an open set  $V_2$  so that  $Q_2 \subset R \subset V_2$  and  $w(V_2 - Q_2) < \varepsilon^2$ . Set now  $Q = Q_1 \cup Q_2$ ,  $V = V_1 \cap V_2$ , then  $Q \subset R \subset V$ ,  $\mu(V - Q) < \varepsilon$  and  $w(V - Q) < \varepsilon^2$ .

Consider now an open cover of  $Q$  consisting of open  $d$ -cells with edges parallel to the coordinate axes and with rational endpoints that are contained in  $V$ . There is a finite subcover whose union  $P$  is such that  $Q \subset P \subset V$  and therefore  $P \Delta R \subset V - Q$  implying  $\mu(P \Delta R) \leq \mu(V - Q) < \varepsilon$  and  $w(P \Delta R) \leq w(V - Q) < \varepsilon^2$ .

It is not hard to see that  $P$  may be partitioned into a finite number of cubes.

(QED claim)

Now let  $G_1, G_2, \dots, G_p$  be the partition of  $P$  into cubes and let  $G_0$  be the smallest cube containing  $R - P$  (the edge of  $G_0$  is not longer than  $d_w$ , the diameter of  $R$ ).

Consider the partitioning of each  $G_i$ ,  $i = 0, 1, 2, \dots, p$  into  $\ell_i^d$  congruent sub-cubes where

$$\begin{aligned} \ell_0^d &\leq K\varepsilon < (\ell_0 + 1)^d \\ \ell_i^d &\leq K(1 - \varepsilon) \frac{\mu(G_i)}{\mu(P)} < (\ell_i + 1)^d \quad i = 1, 2, \dots, p \end{aligned}$$

For large enough  $K$  we have  $\ell_i \geq 1$  for all  $i$ . Assume, then,  $\ell_i \geq 1$  for  $i = 1, 2, \dots, p$ . The diameter of each subcube of  $G_0$  is:

$$\sqrt{d} \frac{d_w}{\ell_0} = \sqrt{d} \frac{(K\varepsilon)^{1/d}}{\ell_0} \frac{d_w}{(K\varepsilon)^{1/d}} < \sqrt{d} \frac{\ell_0 + 1}{\ell_0} \frac{d_w}{(K\varepsilon)^{1/d}} \leq 2\sqrt{d} \frac{d_w}{(K\varepsilon)^{1/d}}$$

The diameter of each of the subcubes in  $G_i$ ,  $i = 1, 2, \dots, p$  is:

$$\sqrt{d} \frac{(\mu(G_i))^{1/d}}{\ell_i} = \sqrt{d} \frac{\ell_i + 1}{\ell_i} \left( \frac{\mu(G_i)}{(\ell_i + 1)^d} \right)^{1/d} > 2\sqrt{d} \left( \frac{\mu(P)}{K(1 - \varepsilon)} \right)^{1/d}$$

The total number of subcubes is  $\underline{K} = \ell_0^d + \sum_{i=1}^p \ell_i^d$  and by our construction  $\underline{K} \leq K\varepsilon + K(1 - \varepsilon) = K$ . Now let  $\sigma/(2\sqrt{d})$  be the uniform upper bound in property (P6), and let  $w_j$  be the restriction of  $w$  to the  $j^{\text{th}}$  subcube  $j = 1, 2, \dots, \underline{K}$ , where  $j = 1, 2, \dots, \ell_0^d$  denote the subcube of  $G_0$  then by properties (P5), (P1), (P3), (P2), (P6) we have

$$\begin{aligned}
 D_K(w) &\leq \underline{D}_K(w) \leq \sum_{j=1}^K D_1(w_j) \leq \sum_{j=1}^{\ell_0^d} \frac{\sigma}{2\sqrt{d}} |w| \cdot 2\sqrt{d} \frac{d_w}{(K\varepsilon)^{1/d}} \\
 &\quad + \sum_{j=\ell_0^d+1}^K \frac{\sigma}{2\sqrt{d}} |w_j| 2\sqrt{d} \left(\frac{\mu(P)}{K(1-\varepsilon)}\right)^{1/d} \\
 &= \sigma\varepsilon^2 |w| \frac{d_w}{(K\varepsilon)^{1/d}} + \sigma |w| \left(\frac{\mu(P)}{K(1-\varepsilon)}\right) < \sigma |w| \left(\frac{|\mu|_w + \varepsilon}{K(1-\varepsilon)}\right)^{1/d} + \sigma \frac{d_w}{K^{1/d}} \varepsilon^{2-1/d}
 \end{aligned}$$

i.e.:

$$\limsup_{K \rightarrow \infty} K^{1/d} D_K(w) < \sigma |w| \left(\frac{|\mu|_w + \varepsilon}{1-\varepsilon}\right)^{1/d} + \sigma d_w \varepsilon^{2-1/d}$$

Letting  $\varepsilon \rightarrow 0$  completes the proof. □

**LEMMA A2:** For any  $\{D_K\}_{K \geq 1}$  that satisfies (P1)-(P7): (i) The lower and upper limits of  $K^{1/d} D_K(w)$  do not depend on the singular part of  $w$ .

(ii) Let  $Q$  be any bounded set containing the support of  $m$ , then the lower and upper limits of  $K^{1/d} D_K(w)$  are continuous in  $m$  over  $L^1(Q)$ .

(In other words if  $\{w_j\}_{j \geq 1}$  is a sequence of demands with supports contained in some common bounded set  $Q$ , so that the corresponding densities  $\{m_j\}_{j \geq 1}$  satisfy  $\lim_{j \rightarrow \infty} \int |m_j - m| d\mu = 0$  where  $m$  is the density of  $w$ , then

$$\lim_{j \rightarrow \infty} (\limsup_{K \rightarrow \infty} K^{1/d} D_K(w_j)) = \limsup_{K \rightarrow \infty} K^{1/d} D_K(w)$$

and

$$\lim_{j \rightarrow \infty} (\liminf_{K \rightarrow \infty} K^{1/d} D_K(w_j)) = \liminf_{K \rightarrow \infty} K^{1/d} D_K(w)$$

**Proof:** Let  $w_1$  and  $w_2$  be two demand distributions with support in  $Q$  and let  $m_1, m_2$  be the densities of their corresponding absolutely continuous parts. Let  $(w_2 - w_1)^+$  be the positive part<sup>9</sup> of  $w_2 - w_1$ , then  $w_2 \leq w_1 +$

---

<sup>9</sup>According to Jordan's decomposition of signed measures.

$(w_2 - w_1)^+$ . Now for  $0 < \varepsilon < 1$  set  $K_2 = [K\varepsilon]$  and  $K_1 = K - K_2$ . By properties (P7), (P1) we have for  $K > 1/\varepsilon$ :

$$D_K(w_2) \leq D_K(w_1 + (w_2 - w_1)^+) \leq D_{K_1}(w_1) + D_{K_2}((w_2 - w_1)^+)$$

Multiplying through by  $K^{1/d}$  and taking to the limit as  $K \rightarrow \infty$  (Noting that  $K_1/K \rightarrow 1 - \varepsilon$  and  $K_2/K \rightarrow \varepsilon$ ) yields by Lemma A1:

$$\begin{aligned} \limsup_{K \rightarrow \infty} K^{1/d} D_K(w_2) &\leq \left(\frac{1}{1 - \varepsilon}\right)^{1/d} \limsup_{K \rightarrow \infty} K^{1/d} D_K(w_1) \\ &\quad + \left(\frac{1}{\varepsilon}\right)^{1/d} \sigma |(w_2 - w_1)^+| \mu_{21}^{1/d} \end{aligned} \quad (A2.1)$$

$$\begin{aligned} \liminf_{K \rightarrow \infty} K^{1/d} D_K(w_2) &\leq \left(\frac{1}{1 - \varepsilon}\right)^{1/d} \liminf_{K \rightarrow \infty} K^{1/d} D_K(w_1) \\ &\quad + \left(\frac{1}{\varepsilon}\right)^{1/d} \sigma |(w_2 - w_1)^+| \mu_{21}^{1/d} \end{aligned} \quad (A2.2)$$

where 
$$\mu_{21} = |\mu|_{(w_2 - w_1)^+} = \mu(S((w_2 - w_1)^+))$$

Noting that  $\varepsilon > 0$  is arbitrary and that the role of indices 1 and 2 in the inequalities (A2.1), (A2.2) can be interchanged, we may deduce the validity of part (i) of the lemma. That is, if  $w_1$  and  $w_2$  differ only in their singular parts (i.e. if  $\mu_{21} = \mu_{12} = 0$ ), then,

$$\limsup_{K \rightarrow \infty} K^{1/d} D_K(w_1) = \limsup_{K \rightarrow \infty} K^{1/d} D_K(w_2)$$

and 
$$\liminf_{K \rightarrow \infty} K^{1/d} D_K(w_1) = \liminf_{K \rightarrow \infty} K^{1/d} D_K(w_2)$$

Having established (i) we may substitute the  $|(w_2 - w_1)^+|$  on the right hand sides of (A2.1), (A2.2) by:



$$\int_{\{x: m_2(x) \geq m_1(x)\}} (m_2(x) - m_1(x)) d\mu(x) \leq \int |m_2 - m_1| d\mu = \|m_2 - m_1\|_{L_1}$$

and consequently the right-most terms of (A2.1), (A2.2) can be replaced by  $(1/\varepsilon)^{1/d} \sigma \|m_2 - m_1\|_{L_1} \mu(Q)^{1/d}$ . (Note that  $\mu_{21} \leq \mu(Q)$ ). Observing the modified inequalities, and noting as before that  $\varepsilon$  is arbitrary and that the indices 1 and 2 may be interchanged we may deduce part (ii) of the Lemma. □

LEMMA A3: For any bounded measurable set  $Q \subset \mathbb{R}^d$  and any  $0 \leq q \leq 1$ ,  $\int m^q d\mu$  is continuous in  $m$  over  $L^1(Q)$  (i.e., if  $\lim_{j \rightarrow \infty} \int |m_j - m| d\mu = 0$  where  $m, m_1, m_2, \dots$  are integrable functions with supports in some bounded set, then  $\lim_{j \rightarrow \infty} \int m_j^q d\mu = \int m^q d\mu$  for all  $0 \leq q \leq 1$ )).

Proof: Note first that by Jensen's inequality

$$\frac{1}{\mu(Q)} \int_Q f^q d\mu \leq \left( \frac{1}{\mu(Q)} \int f d\mu \right)^q \tag{A3.1}$$

Next, by the inequality  $|a + b|^q \leq |a|^q + |b|^q$  that is valid for  $0 \leq q \leq 1$  we have:

$$\int m_1^q d\mu - \int |m_2 - m_1|^q d\mu \leq \int m_2^q d\mu \leq \int m_1^q d\mu + \int |m_2 - m_1|^q d\mu$$

and thus by (A3.1) (substituting  $f = |m_2 - m_1|$ ):

$$|\int m_2^q d\mu - \int m_1^q d\mu| \leq \mu(Q)^{1-q} (\int |m_2 - m_1| d\mu)^q$$

and the assertion clearly follows. □

Definition:  $G_1, G_2, \dots, G_\ell \subset \mathbb{R}^d$  are called separated if

$$\text{Max } \{ \inf \{ \|x - y\| : x \in G_i, y \in G_j \} : 1 \leq i, j < \ell \} > 0$$

Let  $l_{G_i}$  denote the characteristic function of  $G_i$ , i.e.,  $l_{G_i}(x) = 1$  if  $x \in G_i$  and 0 otherwise.

LEMMA A4: The set of functions of the form  $\sum_{i=1}^{\ell} m_i l_{G_i}$ , where  $G_1, G_2, \dots, G_{\ell}$  are separated cubes, is dense in  $L^1(\mathbb{R}^d)$ .

Proof: We know that continuous functions with compact support are dense in  $L^1$  (see [Ru], p. 71, Th. 3.15). Now let  $m$  be a continuous function with the compact support  $R$ .  $m$  is uniformly continuous and bounded on its support  $R$ . For any  $\varepsilon > 0$  there is  $\delta > 0$  so that  $|m(x) - m(y)| < \frac{\varepsilon}{2\mu(R)}$ , whenever  $\|x - y\| < \delta$ . (Note that the case  $\mu(R) = 0$  is trivial since it implies  $m \equiv 0$  in  $L^1$ .) Let  $M = \text{Max}_{x \in R} |m(x)|$ . Using a construction as in the proof of Lemma A1, we may approximate  $R$  by a set  $P$  which is a union of disjoint cubes so that  $\mu(P \Delta R) < \varepsilon/4M$ . If necessary we may further partition the cubes in  $P$  into smaller cubes so as to have  $P = \bigcup_{i=1}^{\ell} G_i$  where the  $G_i$ 's are disjoint cubes with diameters shorter than  $\delta$ . If we wish the cubes to be separated, we may (while keeping their centers fixed) contract each one of them by a  $(1 - \frac{\varepsilon/4M}{\mu(P)})^{1/d}$  scale factor so that the modified union  $\bigcup_{i=1}^{\ell} G_i$  satisfies  $\mu(\bigcup_{i=1}^{\ell} G_i \Delta R) < \varepsilon/2M$ . Now for  $i = 1, 2, \dots, \ell$  let  $m_i$  be some (non-zero) value assumed by  $m$  in  $G_i$  (if  $m(x) \equiv 0$  on  $G_i$ , then delete  $G_i$  from the union).

$$\text{Let } \tilde{m} = \sum_{i=1}^{\ell} m_i l_{G_i} \text{ then:}$$

$$\begin{aligned} \int |m - \tilde{m}| d\mu &= \int_{R \cap (\bigcup_i G_i)} |m - \tilde{m}| d\mu + \int_{R - \bigcup_i G_i} |m - \tilde{m}| d\mu < \frac{\varepsilon}{2\mu(R)} \cdot \mu(R) + M \cdot \frac{\varepsilon}{2M} \\ &= \varepsilon \end{aligned}$$

We have proved, then, that simple functions based on separated cubes are dense in the space of continuous functions with compact support and thus also dense in  $L^1$ . □

LEMMA A5: If  $w$  has a compact support  $S(w)$  and  $\lim_{K \rightarrow \infty} \int f_K dw = 0$ ,  $\{f_K\}_{K \geq 1}$  being equicontinuous, then  $\lim_{K \rightarrow \infty} (\sup_{x \in S(w)} f_K(x)) = 0$ .

Proof: Suppose on the contrary that there is  $\varepsilon > 0$  so that  $\sup_{x \in S(w)} f_K(x) > 2\varepsilon$  for infinitely many  $K$ 's,  $K_1 \leq K_2 \leq K_3 \dots$ . By the equicontinuity assumption there is a  $\delta > 0$  so that  $|f_K(x) - f_K(y)| < \varepsilon$  whenever  $\|x - y\| < \delta$ . And thus for  $K = K_1, K_2, \dots$  there is a ball of radius  $\delta$  in which  $f_K > \varepsilon$ . Now the function  $g(x) = w(B(x, \delta))$ ,  $B(x, \delta)$  being a ball of radius  $\delta$  centered at  $x$  is lower semicontinuous in  $x$  and therefore assumes some minimum  $\alpha > 0$  on  $S(w)$  which is compact. That implies  $\limsup_{K \rightarrow \infty} \int f_K dw \geq \delta \cdot \alpha > 0$  which is a contradiction. □

LEMMA A6: If  $R$  is a bounded measurable subset of  $R^d$  with a boundary of null Lebesgue measure, then for every  $\varepsilon > 0$ , there are sets  $P$  and  $Q$  both finite unions of  $d$ -cells with rational coordinates so that  $P \subset R \subset Q$  and  $\mu(Q - P) < \varepsilon$ .

Proof: Since  $\mu$  is regular there exists an open set  $V$  and a compact set  $T$  so that  $T \subset \text{Int}(R) \subset R \subset \text{cl}(R) \subset V$  and  $\mu(V - \text{cl}(R)) < \varepsilon/2$ ,  $\mu(\text{Int}(R) - T) < \varepsilon/2$  ( $\text{cl}(R)$  and  $\text{Int}(R)$  denote the closure and interior of  $R$ , respectively).

For any point  $x \in T$  consider an open  $d$ -cell with rational endpoints that includes  $x$  but is contained in  $\text{Int}(R)$ . That constitutes an open cover of  $T$ , of which, by compactness of  $T$ , there is a finite subcover. Let  $P$  be the union of such finite subcover  $T \subset P \subset \text{Int}(R)$ . In a similar

fashion (considering the fact that  $\text{cl}(R)$  is compact and contained in the open set  $V$ ) we may construct a finite union of open  $d$ -cells  $Q$  so that  $\text{cl}(R) \subset Q \subset V$ . It follows now that,

$$\begin{aligned} \mu(Q - P) &< \mu(V - T) = \mu(V - \text{cl}(R)) + \mu(\partial R) + \mu(\text{Int } R \cap T) < \varepsilon/2 + 0 + \varepsilon/2 \\ &= \varepsilon \end{aligned}$$

$(\partial R \equiv \text{cl}(R) - \text{Int}(R))$  is the boundary of  $R$ ).

And the proof is complete. □

LEMMA A7: If the supports of  $w_i$  ( $i = 1, 2, \dots, \ell$ ) are separated then there is integer  $M$  with the property that whenever  $K \geq M$ ,  $D_K(\sum_i w_i) = \sum_i D_{K_i}(w_i)$  for some  $K_i \geq 1$  ( $i = 1, 2, \dots, \ell$ ) where  $\sum_i K_i = K$ .

Proof: We first establish the lemma for  $\ell = 2$ . It is enough to show that if  $K$  is sufficiently large no center in the optimal center set  $C_K^*$ , will serve both  $w_1$  and  $w_2$ . Since, by hypothesis,  $\text{Inf} \{ \|x - y\| : x \in S(w_1), y \in S(w_2) \} > 0$  it is enough to prove that:

$$\lim_{K \rightarrow \infty} \left( \sup_{x \in S(w)} \min_{c \in C_K^*} \|x - c\| \right) = 0 \tag{A7.1}$$

To prove (A7.1) note first that:

$$\lim_{K \rightarrow \infty} D_K(w) = \lim_{K \rightarrow \infty} \left( \int \min_{c \in C_K^*} \|x - c\| dw(x) \right) = 0 \tag{A7.2}$$

This follows from the fact that with an unlimited number of centers we can make  $D_K(w)$  arbitrarily small by covering  $S(w)$  with a fine enough grid of center.

Consider now the sequence of functions:

$$f_K(x) \equiv \text{Min}_{c \in C_K^*} \|x - c\| \quad K = 1, 2, \dots$$

Obviously,  $|f_K(x) - f_K(y)| \leq \|x - y\|$  for all  $c, y \in \mathbb{R}^d$  implying that the sequence  $\{f_K\}_{K \geq 1}$  is equicontinuous. From Lemma A5 we know, then, that (A7.2) implies (A7.1)<sup>10</sup> and the proof (for  $\ell = 2$ ) is complete. Repeated application of the lemma with  $\ell = 2$  establishes it for arbitrary  $\ell$ .  $\square$

---

<sup>10</sup>Recall that for general (non-equicontinuous) sequences of functions,  $L^1$  convergence does not imply  $L^\infty$  convergence.

APPENDIX B: Extension of Theorem 1 to the Restricted K-median problem

We prove here that Theorem 1 holds for  $\bar{D}_K(w) \equiv \text{Min} \{D(w, C) : C \subset S(w), |C| \leq K\}$  with the same constant  $\gamma_d$  as for  $D_K(w)$ .

LEMMA B1: For any bounded measurable R

$$\lim_{K \rightarrow \infty} K^{1/d} \bar{D}_K(\mu_R) = \gamma_d (\mu(R))^{(d+1)/d}$$

Remark: Since  $\bar{D}_K$  satisfies all the properties of Lemma 2 except of (P7), we already know (Lemma 3\*) that Lemma B1 holds for R's with a boundary of zero Lebesgue measure.

Proof: For any integer  $n \geq 1$  approximate R by a finite union, P, of d-cells with rational endpoints so that  $\mu(P \Delta R) < 1/n^2$ . Using Lemma A1 and property (P1) of Lemma 2 (note that all the properties in Lemma 2 with the exception of (P7) remain valid for  $\bar{D}_K(w)$ ) we may neglect R - P (just allocate some fixed but small fraction of the K centers to R - P). For the sake of simplicity assume, then, that  $R \subset P$  to begin with.

Since the endpoints of the cells in P are rational, P can be partitioned to a finite number, say  $\ell$ , of congruent cubes. Now for any integer  $K \geq \ell \cdot n$ , set  $m = \left\lfloor \left(\frac{K}{\ell n}\right)^{1/d} \right\rfloor$  and partition each cube into  $m^d$  congruent subcubes. We have now  $\ell m^d$  disjoint congruent cubes  $Q_1, Q_2, \dots, Q_{\ell m^d}$  covering the set R. Locate now n centers (medians) to each subcube  $Q_i, i = 1, 2, \dots, \ell m^d$ , so as to solve the unrestricted n-median problem for the demand  $\mu_{Q_i}$ . Let  $C_i$  be the set of centers for  $Q_i$  (clearly  $C_i \subset Q_i$ ). If  $Q_i \cap R = \emptyset$  let  $\delta_i = 0$ , otherwise let  $\delta_i = \max_{c \in C_i} \inf_{x \in R} \|x - c\|$ , let  $c_i^*$  be some  $c \in C_i$  for which this maximum is

obtained, and let  $T_i$  be the cube of edge  $2\delta_i/\sqrt{d}$  that is centered at  $c_i^*$  (assume same orientation as  $Q_1, Q_2, \dots$ ). Clearly  $Q_i \cap T_i \subset Q_i - R$ . Now since at least one corner point of  $T_i$  as well as its center are contained in  $Q_i$ , it follows that  $Q_i \cap T_i$  and thus also  $Q_i - R$ , contain a  $d$ -orthant of  $T_i$  (i.e., a cube of edge  $\delta_i/\sqrt{d}$ ). Consequently,  $\mu(Q_i - R) \geq d^{-\frac{1}{2}d} \delta_i^d$  and

$$\sum_{i=1}^{\ell m^d} d^{-\frac{1}{2}d} \delta_i^d \leq \mu(P - R) < \frac{1}{n^2} \quad (B1.1)$$

Consider now the feasible solution for the restricted  $K$ -median problem with demand  $\mu_R$ , obtained by substituting each center in  $\bigcup_{i=1}^{\ell m^d} C_i$  by the closest point in the closure of  $R$ .<sup>11</sup> Let  $D_i$  be the cost contributed by the demand in  $Q_i$ . Using Lemma 3 and the triangle inequality we have:

$$D_i \leq (\gamma_d + o(1)) \left( \frac{\mu(P)}{\ell m^d} \right)^{(d+1)/d} n^{-1/d} + \left( \frac{\mu(P)}{\ell m^d} \right) \delta_i$$

where  $o(1)$  is a term tending to 0 as  $n \rightarrow \infty$ .

Consequently:

$$\bar{D}_K(\mu_R) \leq \ell m^d (\gamma_d + o(1)) \left( \frac{\mu(P)}{\ell m^d} \right)^{(d+1)/d} n^{-1/d} + \mu(P) \frac{1}{\ell m^d} \sum_{i=1}^{\ell m^d} \delta_i$$

Noting that  $\frac{1}{\ell m^d} \sum_{i=1}^{\ell m^d} \delta_i \leq \left( \frac{1}{\ell m^d} \sum_{i=1}^{\ell m^d} \delta_i^d \right)^{1/d}$  and (B1.1) it follows that:

---

<sup>11</sup>Note that  $m^d \cdot n \leq K$ , so we have at most  $K$  centers.

$$\begin{aligned} \bar{D}_K(\mu_R) &\leq (\gamma_d + o(1))\mu(P)^{(d+1)/d} (\ell m^d n)^{-1/d} + d^{-1/2d} \mu(P) (\ell m^d n)^{-1/d} n^{-1/d} \\ &= (\gamma_d + o(1))\mu(P)^{(d+1)/d} (\ell m^d n)^{-1/d} \end{aligned}$$

where the equality reflects that  $n^{-1/d} = o(1)$  (as  $n \rightarrow \infty$ ). Noting now that for any fixed  $n$  (and  $\ell$ )  $\lim_{K \rightarrow \infty} \frac{\ell m^d n}{K} = 1$  and that  $\mu(P) = \mu(R) + o(1)$  the proof is complete. □

Corollary B2: Theorem 1 remains valid, with the same constant  $\gamma_d$ , when  $D_K$  is substituted by  $\bar{D}_K$  (in (3.1a, b) and in the definition of  $k_K$ ).

Proof: By Lemma A1 and by subadditivity (P1) we may neglect, for  $K \rightarrow \infty$ , the cost contribution of the singular part of  $w$  as well as of the part of  $w$  supported by subsets where the demand density  $m(x)$  is higher than some arbitrary big number  $M$  (By Lemma A1  $\bar{D}_K(w_{\{x:m(x)>M\}}) = o(K^{-1/d} \mu(\{x: m(x) > M\})^{1/d}) = o(K^{-1/d} M^{-1/d})$ .)

Consider, then, an absolutely continuous demand distribution  $w$  with a bounded density  $m$ . One can approximate such density from above by a density  $m_S$  which is a simple function dominating  $m$  but with the same support as  $m$ . Obviously  $D_K(w) \leq \bar{D}_K(w) \leq \bar{D}_K(w_S)$ , where  $w_S$  denotes the demand associated with  $m_S$ . It is enough, now, to prove the theorem for densities that are simple functions, since if it holds for simple functions, then

$$\begin{aligned} \gamma_d \left( \int m^{d/(d+1)} d\mu \right)^{(d+1)/d} &\leq \liminf_{K \rightarrow \infty} K^{1/d} \bar{D}_K(w) \leq \limsup_{K \rightarrow \infty} K^{1/d} \bar{D}_K(w) \\ &\leq \gamma_d \left( \int m_S^{d/(d+1)} d\mu \right)^{(d+1)/d}, \end{aligned}$$



and if we take a sequence of simple functions  $m_{S_i} \downarrow m$ , then by the bounded convergence theorem  $\int m_{S_i}^{d/(d+1)} d\mu \rightarrow \int m^{d/(d+1)} d\mu$ , establishing convergence for  $K^{1/d} \bar{D}_K(w)$ . Assume, then, that  $m$  is simple, that is,  $w =$

$\sum_j m_j \mu_{R_j}$ . Let  $K_i = \left[ \frac{m_i^{d/(d+1)} \mu_{R_i}}{\sum_j m_j^{d/(d+1)} \mu_{R_j}} K \right]$ . By subadditivity (P1),  $\bar{D}_K(w) \leq \sum_i \bar{D}_{K_i}(m_i \mu_{R_i})$  and thus using Lemma B1 and Theorem 1 we have after some rearrangement:

$$\limsup_{K \rightarrow \infty} K^{1/d} \bar{D}_K(w) \leq \gamma_d (\sum_i m_i^{d/(d+1)} \mu_{R_i})^{(d+1)/d} = \lim_{K \rightarrow \infty} K^{1/d} D_K(w)$$

And recalling that  $\bar{D}_K(w) \geq D_K(w)$  we conclude that  $K^{1/d} \bar{D}_K(w)$  converges to the same limit as  $K^{1/d} D_K(w)$ . It can be easily shown now that the asymptotic center allocation for the restricted  $K$ -median problem is the same as for the unrestricted problem, completing the proof of Theorem 1 for the restricted  $K$ -median problem. □

APPENDIX C: Optimal Branching Bifurcation Angles in Euclidean Networks

Suppose we have a finite set of sources and sinks in the plane (with some of the sources possibly without specified supply and/or location) and suppose that transporting of units along an arc of length  $\ell$  costs  $c(f)\ell$ . What is the geometry of the optimal network (of roads, pipelines, cables, etc.) supporting the flow requirement?

In the case where  $c(f)$  is linearly increasing in  $|f|$ <sup>12</sup> (or convex with  $c(0) = 0$ ), it is clear that in the optimal network there will be flow only along straight line segments joining sources and sinks (i.e., we will not have any transshipment nodes). In other cases, however, we may have (essentially because of fixed costs and economies of scale) junctions (nodes) that are neither sources nor sinks.

In the case where  $c(f)$  is a positive constant (i.e., cost is independent of the flow) one may readily see that the optimal network will consist of a forest of Steiner trees, each component being a network of shortest total length joining a balanced set of sources and sinks. Arcs may join at points other than sources or sinks. The angles the arcs adjacent at such transshipment node of degree 3 are known to be  $120^\circ$  (see [Co] for example).

This well-known result may be generalized in a way that is parallel to the generalization of the solution of the 1-median problem for 3 identical customers (Fermat Problem, see for example [K2]) to the

---

<sup>12</sup>It is assumed here and in the rest of the discussion that  $c(f) = c(-f)$  for all  $f$ .

solution of the 1-median problem with three weighted customers (Weber Problem [We], or the generalized Fermat Problem [K2]).<sup>13</sup>

Consider, then, a transshipment node, or a source node whose location was not pre-specified in the problem, which we call a "free" node, of degree 3 as in Figure C1 below.

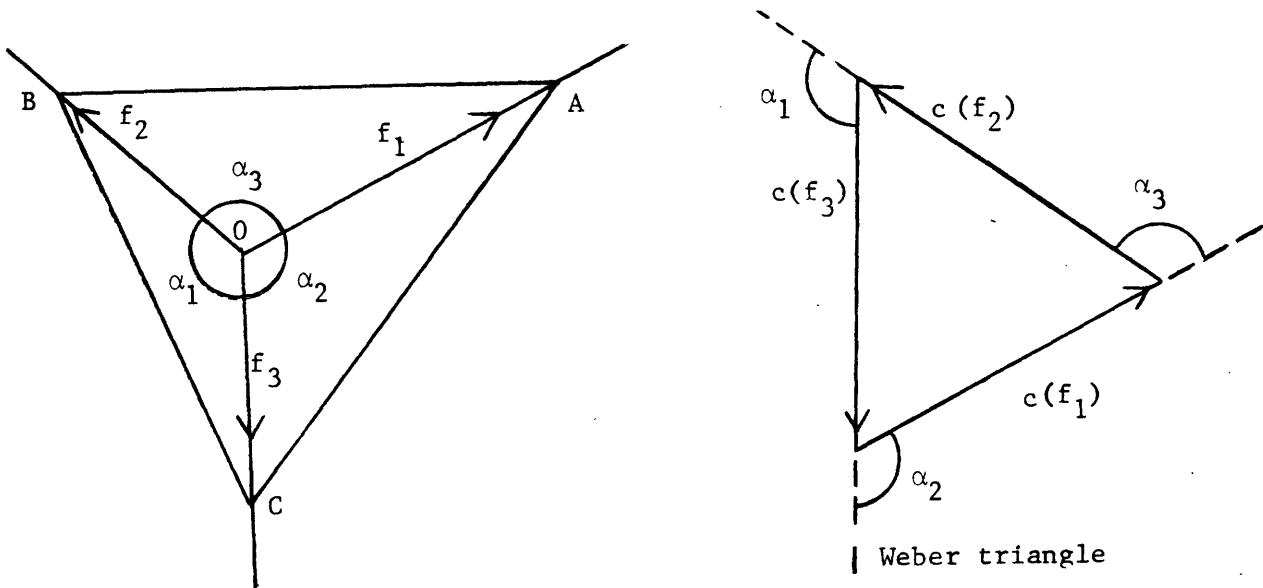


FIGURE C1: Angles in a "free" node of degree 3.

Fix any 3 points A, B and C on the adjacent arcs carrying the flows  $f_1$ ,  $f_2$  and  $f_3$  out of the "free" node 0. The point 0, obviously minimizes

$$c(f_1) \cdot |OA| + c(f_2) \cdot |OB| + c(f_3) \cdot |OC|$$

<sup>13</sup>Courant and Robbins [Co] consider the problem of finding the point in the plane with the smallest sum of distances to  $n$  given points for  $n > 3$  to be a "sterile" generalization of the problem with  $n = 3$ , and view finding the shortest connected network (tree) joining the  $n$  points as "the natural" generalization. Kuhn [K2] on the other hand strongly contests their statement. The network design problem formulated here seems to include both generalizations and thus brings the peace to that debate.

that is it solves the weighted 1-median (i.e. general Fermat, or Weber) problem with weights  $c(f_1)$ ,  $c(f_2)$  and  $c(f_3)$  at A, B and C. The gradient of the cost with respect to the position of 0 is

$$c(f_1) \frac{\vec{OA}}{|OA|} + c(f_2) \frac{\vec{OB}}{|OB|} + c(f_3) \frac{\vec{OC}}{|OC|}$$

( $\vec{OA}$  denotes the displacement from 0 to A, and  $|OA|$  denotes the length of that displacement) and thus should be null if 0 is a "free" node. Consequently by the triangle law of vector addition the angle  $\alpha_1$  between  $f_j$  and  $f_k$  is the external angle at the vertex opposite  $c(f_1)$  in a triangle of edges  $c(f_1)$ ,  $c(f_2)$  and  $c(f_3)$ . This triangle is known in the context of the 1-median problem as Weber triangle. By the law of sines we know that:

$$\frac{\sin \alpha_1}{c(f_1)} = \frac{\sin \alpha_2}{c(f_2)} = \frac{\sin \alpha_3}{c(f_3)} \quad (C.1)$$

If  $c(f)$  is homogenous in  $|f|$  i.e.,  $c(f) = b|f|^\beta$  for some  $\beta \geq 0$  and  $b > 0$ , then the angles  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_3$  solving (C.1) depend only on the ratios  $f_2/f_1$ ,  $f_3/f_1$ .

Examples:

(a) If  $f_1 = f_2 = f_3$  (or if  $\beta = 0$ ) we have  $\sin \alpha_1 = \sin \alpha_2 = \sin \alpha_3$  and thus (recalling  $\alpha_1 + \alpha_2 + \alpha_3 = 360^\circ$ ) we have  $\alpha_1 = \alpha_2 = \alpha_3 = 120^\circ$ .

(b) If we have a transshipment node with  $f_1 + f_2 + f_3 = 0$  and  $f_2 = f_3$  and if  $\beta = \frac{1}{2}$  then  $(\sin \alpha_1)/\sqrt{2} = \sin \alpha_2 = \sin \alpha_3$  which is solved by  $\alpha_1 = 90^\circ$  and  $\alpha_2 = \alpha_3 = 135^\circ$ .

Figure C2 below demonstrates the solution of a single facility (source) location problem for a given set of identical customers (sinks) given that  $c(f) = b|f|^{\frac{1}{2}}$ . The geometry is such that in the solution all the free nodes are of the type  $f_1 = f_2 = f_3$  or transshipment nodes ( $f_1 + f_2 + f_3 = 0$ ) with  $f_2 = f_3$  as in examples (a) and (b) above.

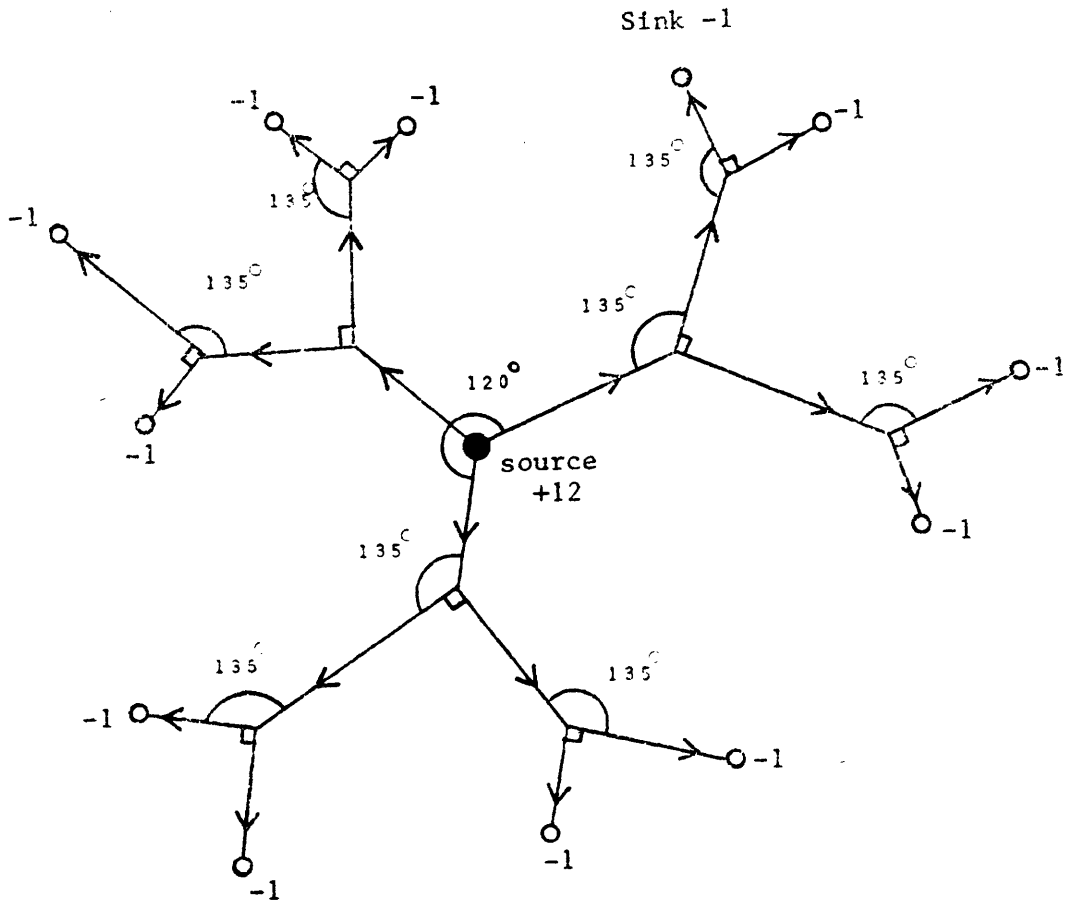


FIGURE C2: Solution of a Euclidean network flow problem with  $c(f) = b\sqrt{|f|}$  and one "free" source.

APPENDIX D: Calculation of  $D_K(w)$  for Cantor Distribution

Note first that if the demand in the subinterval  $[0, 1/3]$  (or  $[1/3, 1]$ ) is served by a single center outside the subinterval, then the associated cost is equal to the product of  $1/2$  (the mass in the subinterval) and the distance between the center and the middle of the subinterval. If, on the other hand, the center lies inside the subinterval the associated cost will be higher than the associated product. Consequently, if we have a single center  $K = 1$  we should place it anywhere in the middle subinterval  $[1/3, 2/3]$  and the optimal cost  $D_1(w)$  will be:

$$D_1(w) = \frac{1}{2} \left( \begin{array}{l} \text{mass in each of the sub-} \\ \text{intervals } [0, 1/3] \text{ and } \\ [2/3, 1] \end{array} \right) \cdot \frac{2}{3} \left( \begin{array}{l} \text{the distance between the} \\ \text{middles of } [0, 1/3] \text{ and } \\ [2/3, 1] \end{array} \right)$$

Observe now that if we have two centers or more ( $K \geq 2$ ), then the subintervals  $[0, 1/3]$  and  $[2/3, 1]$  will not share any common center. Applying this logic recursively we conclude that any optimal solution of the  $K$ -median problem partitions Cantor set into  $K$  subsets each being a scaled-down copy of Cantor set served by a single center lying somewhere on its middle subinterval. The cost associated with such single center subset is of the form  $(1/3)6^{-j} = (1/3)(1/2^j)(1/3^j)$  for some  $j \geq 0$  ( $1/3$  is the cost for the original scale Cantor set with one center.  $1/2^j$  and  $1/3^j$  are the demand and the scale, scaling factors.) Consequently:

$$D_K(w) = \text{Min}_{j_1, j_2, \dots, j_k} \left\{ \sum_{i=1}^K \frac{1}{3} 6^{-j_i} : \sum_{i=1}^K 2^{-j_i} = 1 \right\} \quad (D.1)$$

Let  $\ell = \lfloor \log_2 K \rfloor$ , i.e.,  $2^\ell \leq K < 2^{\ell+1}$ .

LEMMA D1:  $j_1, \dots, j_k$  are optimal in (D.1) if and only if  $2^{\ell+1} - K$  of the  $j_i$ 's are equal to  $\ell$  and the rest of the  $j_i$ 's are equal to  $\ell + 1$ .

Proof: Let  $j_1, j_2, \dots, j_k$  be feasible in (D.1), and let  $t = \max(j_1, \dots, j_k)$ . The number of  $j_i$ 's with  $j_i = t$  is even. To see this, note first that if  $j_i = t$  for all  $i = 1, 2, \dots, k$ , then  $K = 2^t$  which is even, otherwise the sum  $\sum_{\{i: j_i < t\}} 2^{-j_i}$  is a multiple of  $2^{-(t-1)}$  and there-

fore its complement (to 1)  $\sum_{\{1: j_i = t\}} 2^{-j_i} = |\{i: j_i = t\}| \cdot 2^{-t}$  should be a multiple of  $2^{-(t-1)}$  too, i.e.,  $|\{i: j_i = t\}|$  is even. We shall show now that if  $j_1, \dots, j_k$  is optimal, then for all  $i = 1, 2, \dots, k$   $j_i \geq t - 1$ . Suppose on the contrary that there is an  $i$  so that  $j_i = t - s$  with  $s \geq 2$ . Observing equality,

$$2 \cdot 2^{-t} + 2^{-(t-s)} = 2^{-(t-1)} + 2 \cdot 2^{-(t-s+1)}$$

and the associated inequality,

$$2 \cdot 6^{-t} + 6^{-(t-s)} > 6^{-(t-1)} + 2 \cdot 6^{-(t-s+1)}$$

(Note that since by hypothesis  $s \geq 2$ , we have  $t > t - s + 1$  and thus  $6^{-t} < 6^{-(t-s+1)}$  and hence  $2(6^{-t}) + 6(6^{-(t-s+1)}) > 6(6^{-t}) + 2(6^{-(t-s+1)})$ ), we deduce that substituting a pair of the  $j_i$ 's that are maximal ( $= t$ ) and a single  $j_i$  that is smaller than  $t - 1$  (i.e.  $= t - s$ ) by a pair of  $j_i$ 's that are equal to  $t - s - 1$  and a single  $j_i$  that is equal to  $t - 1$  will reduce the cost, contradicting optimality. Thus for all  $i$   $t - 1 \leq j_i \leq t$ . It is not hard to see now that the only feasible solution satisfying such a condition is the one characterized in the assertion of the

lemma. It is not hard to see now that an optimal solution of (D.1) should exist and that the characterization in the statement is sufficient (since the  $j_i$ 's are permutable). □

Corollary D2: If  $w$  is the Cantor distribution then,

$$D_K(w) = \frac{1}{3} \left( \frac{K}{2^\ell} \right)^{\log_2 3} \left( \frac{5}{3} - \frac{2}{3} \frac{K}{2^\ell} \right) K^{-\log_2 3}$$

where  $\ell = \lceil \log_2 K \rceil$ .