154

INDUSTRIAL LOCATION ON THE EUCLIDEAN PLANE

Arnold M. Faden*

1. INTRODUCTION

One of the most important problems in location theory is that of arranging a system of facilities to meet a spatially-distributed demand in an optimal manner. The most familiar example is the location of manufacturing plants, but the problem arises also for grain elevators, retail stores, schools, hospitals, police and fire stations and many other facilities.[1]

An arrangement may be characterized abstractly as follows: (a) A set of points is chosen at which "plants" are to be located (the set may be finite or infinite in number). (b) A pattern of transportation flows is specified, from plants to surrounding territory. The total outflow from each plant equals its production level.

We assume the following data are available: (1) For each point of space, a cost function, representing the cost incurred in producing that level of output in a plant located at that point. (2) A transportation cost function, defined over all point-pairs and shipment sizes.

At this point we may distinguish two versions of the optimal location problem. The more restrictive version specifies: (3a) a delivery

---

* The author is Assistant Professor of Economics, Iowa State University, Ames, Iowa.

1 For an overview of such "service systems" see Faden [6].

requirement measure over space (for example, in the form of a density function). One is then to minimize the sum total of production and transportation costs while satisfying those required deliveries. The more general version specifies: (3b) for each point of space, a benefit function depending on deliveries to that point. One is then to maximize the sum total of benefits minus production costs minus transportation costs. We shall be concerned mainly with the first of these formulations.

Two broad approaches to these problems may be distinguished in the literature. The first, and generally more recent, is to devise efficient algorithms for finding optimal, or near-optimal solutions.[2] The second is to make further idealizing assumptions, and derive the solution in general mathematical form. This approach, which may be called the Christaller-Lösch tradition ([2], [12]), is the one we pursue here.

Let us, then, strengthen assumptions (1), (2), and (3a) or (3b) as follows: (1') Production cost conditions are uniform; that is, the cost functions are the same for all points of space. (2') Space itself is the ordinary Euclidean plane of high-school geometry, with its ordinary metric; transportation cost is proportional to distance, and to quantity shipped. (3a') For the restrictive version, delivery requirements are uniform over the entire plane.[3] (3b') For the general version, benefit conditions are

---

[2] For example, Cooper [3], Kuehn and Hamburger [11].

[3] This assumption is more realistic than it seems at first glance, for many systems. For example, law or custom may impose uniform servicing requirements for police, fire, street-cleaning, or school systems. The custom of freight absorption over short distances may lead to uniform demands in the case of water supply, electricity, gas and sewage systems. For further discussion of this point see Faden [6].

uniform; that is, benefit, as a function of delivery density, is independent of position on the plane.

Under assumptions (1'), (2'), and (3a') or (3b'), the Christaller-Lösch tradition comes up with some rather strong conclusions concerning the optimal arrangement: (a') The plant locations form a *honeycomb lattice*. (b') Each plant is the exclusive supplier of its *dirichlet region*.[4]

Since most of our further discussion will revolve around these concepts, let us define them formally.[5] A *lattice* (in two dimensions) is a set of points of the form mX + nY, where X and Y are fixed, linearly-independent vectors, and *m* and *n* range independently over the integers -- positive, negative, and zero. The origin of coordinates may be chosen arbitrarily on the plane. Thus in Figure 1, nine points of a lattice are shown, corresponding to m, n = -1, 0, 1. (All of these except the origin are circled.) A *honeycomb lattice* is a lattice in which the defining vectors X and Y are of equal length, and the angle between them ($\vartheta + \beta$ in Figure 1) is 60°. The *dirichlet region* of a lattice point is that portion of the plane closer to it than to any other lattice point. In Figure 1, the dirichlet region of the origin is the irregular hexagon (abcdef), and the borders of adjacent dirichlet regions are shown extending from these six points. One easily shows that all dirichlet regions for a

---

4 We do not claim that assumptions (1'), (2'), and (3a') or (3b'), and conclusions (a') and (b'), are stated explicitly by all writers in this tradition, but merely that they all seem to have this model in mind, with minor modifications. (In particular, our benefit functions are usually represented as demand functions.)

5 See also Cassels [1; Chap. I], Coxeter [4; pp. 50ff], Dacey [5], and Fejes-Tóth [7].

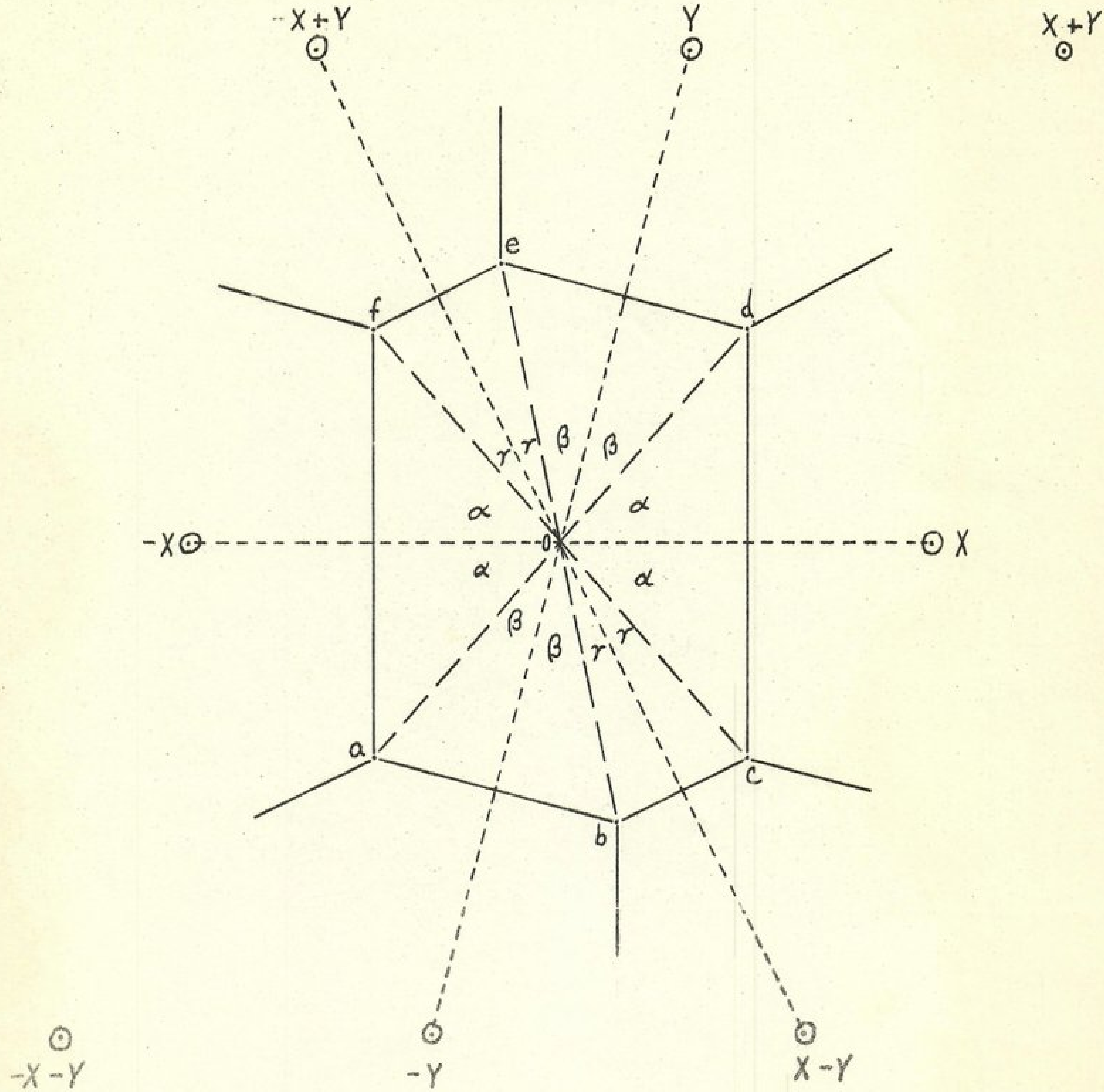

Figure 1. LATTICE WITH DIRICHLET REGIONS

given lattice are congruent. For the honeycomb lattice only, they are regular hexagons. A partition of space into dirichlet regions may be defined in the same way for any arbitrary collection of isolated points.

What is the evidence for the assertion that (a') and (b') characterize the optimal arrangement? An examination of the arguments[6] reveals that it is based more on faith than on demonstration. The more careful arguments([9], [13]) compare the three possible ways of covering the plane by dirichlet regions which are congruent regular polygons, viz, regular hexagons, squares, and equilateral triangles, and, indeed, the regular hexagon -- corresponding to the honeycomb lattice -- proves to be the best of the three. But this is only three cases out of an infinite number of possibilities. How do we know that some irregular arrangement of plants will not do better than any of these? Is it really optimal for plants to supply just their own dirichlet regions? Should all plants produce the same output? Furthermore, do the answers to these questions depend on the form of the production cost function, or the benefit function?

The problem is, in fact, one of extreme difficulty, and the answers to these questions are unknown at the present time.

The major aim of this paper is to prove that the Christaller-Lösch solution is indeed optimal in a certain class of possible arrangements, viz, the class of uniform required deliveries, with plants arranged in a lattice, and each plant the exclusive supplier of its dirichlet region. That is, we restrict our attention to arrangements satisfying condition (b')

---

6 Christaller [2; p. 63f], Haggett [8; p. 49], Isard [9; pp. 240-242], Lösch [12; pp. 110-114], Mills and Lav [13].

of the Christaller-Lösch solution and prove that the honeycomb lattice is optimal in the class of all lattices. Essentially no restrictions are placed on the production cost function. This class, though still a modest subset of the class of all possible arrangements, is much more extensive than any dealt with so far.

Perhaps the most interesting aspect of this result is that even in this narrow class the proof required is so tedious and complicated.[7] This gives a measure of the difficulty in store for one trying to prove the general case.

One last preliminary difficulty remains to be resolved. We are to minimize the sum total of production and transportation costs (or maximize total benefits minus costs, in the general version). If taken literally, this criterion becomes useless, since, in general, any arrangement meeting delivery requirements will entail infinite costs over the whole plane.

For the class of arrangements under consideration here, a simple extension of the original criterion suggests itself. In each arrangement, the plane is covered by an infinite repetition of identical dirichlet regions, with identical production and shipment patterns within each region, and no shipments across borders. Under these circumstances, it seems reasonable to <u>minimize the ratio of total production plus transportation costs in a dirichlet region to the area of a dirichlet region</u>. This will be called the <u>minimal cost-density criterion</u>. It appears to be the criterion used, implicitly or explicitly, by previous studies in the

---

7 We will withdraw this statement in admiration if someone should find, say, a twenty-line proof of our main theorem.

the dirichlet regions degenerate into rectangles. If all three angles are positive the dirichlet regions are hexagons. The honeycomb lattice is exactly the case in which $\alpha = \beta = \gamma = 30°$.

Any generating vectors X, Y determine R, α, β, γ. Conversely, any $R > 0$, and any $\alpha, \beta, \gamma \geq 0$, with at least two angles positive, and satisfying (2), determine a lattice. R (or A, for that matter) may be thought of as determining the "size" or "spacing" of the lattice, while α, β, γ determine its "shape". These are more convenient parameters for us to use than the original X and Y.

The dirichlet area, A, is easily found in terms of R, α, β, γ by summing over the six isosceles triangles:

(3) $$A = R^2(\sin 2\alpha + \sin 2\beta + \sin 2\gamma).$$

Total transportation cost for the dirichlet region is

(4) $$\frac{4}{3}\rho R^3\left[\cos^3\alpha\int_0^{\alpha} \sec^3\theta d\theta + \text{similar terms in } \beta \text{ and } \gamma\right].$$

(This is most easily computed by considering each of the twelve right triangles separately, integrating over a sliver from the origin to the edge of the region, integrating over the central angle, and summing over the triangles.)

By the cost-density criterion, we are to minimize

(5) $$\frac{1}{A}\left[C(P) + \text{total transportation cost}\right],$$

the latter term being given by (4).

B. Let us use (1) and (3) to solve for A and R in terms of P, α, β, γ, and substitute these results in (5). We obtain, after simplification,

$$(6) \qquad \rho C(P)/P + \frac{4}{3}\sqrt{\rho P}\left[\frac{\cos^3\alpha\int_0^\alpha \sec^3 \theta d\theta + \beta \text{ and } \gamma \text{ terms}}{(\sin 2\alpha + \sin 2\beta + \sin 2\gamma)^{3/2}}\right]$$

as the objective function to be minimized.

We have chosen as our lattice parameters, P, α, β, γ. The three angles α, β, γ may range freely, subject to (2) and to the non-negativity and positivity conditions already mentioned, and this range of choice is independent of the value of P chosen. Assumptions (3) and (5) guarantee that there is an optimal P* which is positive.

These observations, together with the form of (6), make it obvious that the optimal values α^*, β^*, γ^* do not depend at all on P*, but should be chosen to minimize

$$(7) \qquad \frac{\cos^3\alpha\int_0^\alpha \sec^3 \theta d\theta + \cos^3\beta\int_0^\beta \sec^3 \theta d\theta + \cos^3\gamma\int_0^\gamma \sec^3 \theta d\theta}{(\sin 2\alpha + \sin 2\beta + \sin 2\gamma)^{3/2}}$$

subject to (2) and to α, β, $\gamma \geq 0$, and to not more than one of those equalling zero.

Thus the optimal "shape" of the lattice may be determined independently of its "size", and is independent of the level of required deliveries, ρ, and of the form of the production cost function (except that these must satisfy assumptions (3) and (5)).

The rest of this proof is directed to showing that the formidable expression (7) has a unique minimum at $\alpha = \beta = \gamma = 30°$ -- the honeycomb lattice case.

C. The denominator of the objective function (7) is positive for all feasible values of α, β, γ. Hence a <u>necessary</u> condition that a given set of values minimize (7) is that they minimize the numerator subject to the value of the denominator being fixed. This suggests looking at the following problem:

Minimize

(8) $$\cos^3\alpha \int_0^{\alpha} \sec^3 \theta d\theta + \cos^3\beta \int_0^{\beta} \sec^3 \theta d\theta + \cos^3\gamma \int_0^{\gamma} \sec^3 \theta d\theta$$

subject to α, β, $\gamma \geq 0$, and to

(9) $$\alpha + \beta + \gamma = 90^\circ, \quad \text{and}$$

(10) $$\sin 2\alpha + \sin 2\beta + \sin 2\gamma = K.$$

Here (8) is the numerator of (7), (9) is the same as (2), and (10) is an auxiliary constraint introduced to fix the value of the denominator of (7) at $K^{3/2}$, K being a parameter. (8) to (10) will be called the <u>angle problem</u>.

The minimal value of the objective function (8) of the angle problem depends on K. Let us write it as $\varphi(K)$. In terms of K, (7) becomes

(11) $$\varphi(K)\, K^{-3/2}.$$

Optimal K* is the one minimizing the expression (11). The optimal angles α^*, β^*, γ^* will be the ones minimizing (8) for this value of K inserted in (10).

Let us now determine the range of values of K which give us even a feasible solution to the angle problem. Zero is clearly the lower limit, though not itself a permissible K value. To find the upper limit, note that, given (9), $\sin 2\alpha + \sin 2\beta + \sin 2\gamma$ may be rewritten as $\sin 2\alpha + 2 \cos\alpha \cos(\beta - \gamma)$ (since $\sin 2\beta + \sin 2\gamma = \cos(2\beta - 90^\circ) + \cos(90^\circ - 2\gamma) = \cos((\beta-\gamma) - \alpha) + \cos((\beta-\gamma) + \alpha) = 2\cos\alpha\cos(\beta-\gamma)$). Thus, for a given value of α, K attains its maximum when $\beta = \gamma$.

By symmetry, K attains its upper limit when $\alpha = \beta = \gamma = 30^{\circ}$ (and the upper limit is therefore $3 \sin 60^{\circ} = 3\sqrt{3}/2$). Conversely, if K is at its upper limit, the only possible values satisfying the conditions of the angle problem are $\alpha = \beta = \gamma = 30^{\circ}$.

These observations furnish the key to our method of proof. Suppose it could be shown that (11) decreases as K increases. Since we are minimizing, K should then be taken as large as possible -- i.e., $K^* = 3\sqrt{3}/2$, and for this value $\alpha^* = \beta^* = \gamma^* = 30^{\circ}$, and we will have established the optimality of the honeycomb lattice.

Thus we have reduced the problem to showing that $\varphi(K)K^{-3/2}$ is a decreasing function of K. It will suffice to prove this for K in the <u>open</u> interval $(0, 3\sqrt{3}/2)$, since the function is continuous at $K = 3\sqrt{3}/2$.

D. Let us form the Lagrangean for the angle problem:

$$(12) \quad \cos^3\alpha \int_0^{\alpha} \sec^3\theta\, d\theta + \beta \text{ and } \gamma \text{ terms} - \lambda(\alpha + \beta + \gamma) - \mu(\sin 2\alpha + \sin 2\beta + \sin 2\gamma).$$

Write $\alpha_k, \beta_k, \gamma_k$, for optimal angle values for a given K. If $\alpha_k > 0$, the derivative of (12) with respect to α must equal zero at these optimal values. Thus

$$(13) \quad 1-3\cos^2\alpha_k \sin\alpha_k \int_0^{\alpha_k} \sec^3\theta\, d\theta - \lambda - 2\mu\cos 2\alpha_k = 0$$

if $\alpha_k > 0$, with similar expressions in β_k and γ_k, provided they are positive.

Formula (13) may be rewritten in the more convenient form

$$(14) \quad 4\mu\sin^2\alpha_k - 3\cos^2\alpha_k \sin\alpha_k \int_0^{\alpha_k} \sec^3\theta\, d\theta = 2\mu + \lambda - 1,$$ using the identity: $\cos 2\alpha_k = 1-2\sin^2\alpha_k$.

If $\alpha_k, \beta_k, \gamma_k > 0$, multiply the α, β, and γ equations (14) by $\cot\alpha_k$, $\cot\beta_k$, $\cot\gamma_k$, respectively, and add. We obtain let us take $\alpha_k \neq \beta_k$. Consider a slightly modified version of the angle problem, (8) - (10), in which both K and γ are taken as parameters. (9) and (10) then determine a pair of values to be assigned to α and β (in any order), and substitution of these values determines (8) uniquely. Let us write the resulting value of (8) as $h(K, \gamma)$. Between h and φ we have the relation

$$\varphi(K) = \min_{\gamma} h(K, \gamma) = h(K, \gamma_k). \tag{20}$$

Differentiating, we get

$$\varphi'(K) = \frac{\partial}{\partial K} h(K, \gamma_k) + \frac{\partial}{\partial \gamma_k} h(K, \gamma_k) \frac{d\gamma_k}{dK} = \frac{\partial}{\partial K} h(K, \gamma_k), \tag{21}$$

the last equality following from the fact that γ_k is optimal. From (8) we obtain

$$\frac{\partial}{\partial K} h(K, \gamma_k) = \frac{d}{d\alpha_k} (\cos^3\alpha_k \int_0^{\alpha_K} \sec^3 \theta d\theta) \frac{\partial \alpha_k}{\partial K} + \frac{d}{d\beta_k} (\cos^3\beta_k \int_0^{\beta_K} \sec^3 \theta d\theta) \frac{\partial \beta_k}{\partial K} \tag{22}$$

where γ_k is held fixed in the evaluation of $\frac{\partial \alpha_k}{\partial K}$ and $\frac{\partial \beta_k}{\partial K}$.

From (9) we obtain

$$\frac{\partial \alpha_k}{\partial K} + \frac{\partial \beta_k}{\partial K} = 0, \tag{23}$$

and from (10) we obtain

$$2 \cos 2\alpha_k \frac{\partial \alpha_k}{\partial K} + 2 \cos 2\beta_k \frac{\partial \beta_k}{\partial K} = 1. \tag{24}$$

(23) and (24) may be solved simultaneously for $\frac{\partial \alpha_k}{\partial K}$ and $\frac{\partial \beta_k}{\partial K}$:

$$\frac{\partial \alpha_k}{\partial K} = - \frac{\partial \beta_k}{\partial K} = \frac{1}{2(\cos 2\alpha_k - \cos 2\beta_k)} = \frac{1}{4(\cos^2\alpha_k - \cos^2\beta_k)}. \tag{25}$$

If these results are substituted into (22), the differentiations carried out there, and the results referred back to (21), we obtain

$$(26)\quad 4\varphi'(K) = \frac{3\cos^2\beta_k \sin\beta_k \int_0^{\beta_k} \sec^3\theta d\theta - 3\cos^2\alpha_k \sin\alpha_k \int_0^{\alpha_k} \sec^3\theta d\theta}{\cos^2\alpha_k - \cos^2\beta_k}$$

which is the expression we were seeking.

F. Since $\alpha_k \neq \beta_k$, at least one of them, say α_k, is positive. Thus formula (19) obtains. Let us now substitute the expression for φ' in (26) into the right-hand side of (19).

After the smoke clears, we are left with the following expression:

$$(27)\quad \frac{3\sin^2\alpha_k \sin^2\beta_k}{\cos^2\alpha_k - \cos^2\beta_k}\left[\frac{\cos^2\beta_k}{\sin\beta_k}\int_0^{\beta_k}\sec^3\theta d\theta - \frac{\cos^2\alpha_k}{\sin\alpha_k}\int_0^{\alpha_k}\sec^3\theta d\theta\right].$$

The sign of (27) is the same as the sign of $\frac{d}{dK}(\varphi K^{-3/2})$, by (19). (27) is symmetrical in α_k and β_k, so without loss of generality we may assume $\alpha_k > \beta_k$.

Outside the bracket, the numerator is non-negative, and positive if β_k is positive. The denominator is negative. (27) would have the negative sign we are seeking if it could be shown that the bracketed expression were positive, and this in turn would be established if we could show that the function

$$(28)\quad f(x) = \frac{\cos^2 x}{\sin x}\int_0^{x}\sec^3\theta d\theta$$

were <u>decreasing</u> in the relevant range (the open interval 0° - 90°). To this final task we now turn.

G. Using the fact that $\int_0^x \sec^3 \theta \, d\theta = 1/2 \sec x \tan x + 1/2 \log (\sec x + \tan x)$, we find from (28) that

$$\text{(29)} \quad \frac{df}{dx} = 1/2 \cot x - \left(\cos x + \frac{\cos^3 x}{2 \sin^2 x} \right) \log (\sec x + \tan x).$$

We must show that this is negative in $(0^\circ, 90^\circ)$.

Start from the obvious inequalities

$$\text{(30)} \quad \sec \theta > 1 > \frac{\cos^3 \theta}{(\sin^2 \theta + 1)^2} \quad ,$$

valid for $0^\circ < \theta < 90^\circ$.

Integrating (30), we obtain

$$\text{(31)} \quad \int_0^x \sec \theta \, d\theta > \int_0^x \frac{\cos^3 \theta \, d\theta}{(\sin^2 \theta + 1)^2} \quad , \text{ or}$$

$$\text{(32)} \quad \log (\sec x + \tan x) > \frac{\sin x}{\sin^2 x + 1} \qquad (0^\circ < x < 90^\circ)$$

Multiply both sides by the positive number $\left(\cos x + \frac{\cos^3 x}{2 \sin^2 x} \right)$:

$$\text{(33)} \quad \left(\cos x + \frac{\cos^3 x}{2 \sin^2 x} \right) \log (\sec x + \tan x) > \frac{\sin x}{\sin^2 x + 1} \left(\cos x + \frac{\cox^3 x}{2 \sin^2 x} \right) =$$

$$\frac{2 \cos x \sin^2 x + \cos^3 x}{2 (\sin^2 x + 1) \sin x} = \frac{1}{2} \left(\frac{2 \cos x \sin^2 x + \cos^3 x}{2 \sin^2 x + \cos^2 x} \right) \frac{1}{\sin x}$$

$$= \frac{\cos x}{2 \sin x} = \frac{1}{2} \cot x.$$

Comparison with (29) shows that f (x) is decreasing.

H. This last result proves that (27) is non-positive for all K in the range $(0, 3\sqrt{3}/2)$, and in fact negative if α_k and β_k are both positive. However, for $K > 2$, it follows from (10) that __all__ feasible angles must be positive (since the sum of two sines cannot exceed 2). Therefore, (27) is actually negative in a neighborhood of $K = 3\sqrt{3}/2$. It follows that $\varphi K^{-3/2}$ is uniquely minimized at $K^* = 3\sqrt{3}/2$. Therefore $\alpha^* = \beta^* = \gamma^* = 30^\circ$, and the optimality of the honeycomb lattice is established. The proof is complete.

3. THE PROBLEM OF INFINITE COSTS[9]

We used cost-density minimization as our criterion in the previous section. This was possible because the class of arrangements under discussion all had the property of covering the plane with an infinite repetition of identical production and transportation flows. But we want to include the possibilities of having plants arranged irregularly, or having plants supply other than their own dirichlet regions. In these cases cost-density is not even defined, there being no "typical" regions over which to minimize the ratio of cost to area. So if we are ever to find the best of all possible arrangements, we will first have to find a stronger criterion than cost-density.

This difficulty may be overcome in a fairly satisfactory manner, but before doing so let us answer one plausible objection which may occur to the practical-minded reader. Why bother with an infinite plane at all?

---

[9] While the discussion in this section is phrased in terms of cost minimization, it applies almost verbatim to the more general criterion of maximizing benefits minus costs.

After all, the entire Earth is finite, so that using a finite space is more realistic, and also avoids the difficulty just mentioned.

The answer lies in the simplicity achieved through the approximation of an unbounded plane. However difficult the optimal location problem is for the plane, it is probably worse for, say, a spherical surface, or a bounded region. The use of infinite approximations to finite situations has often been accompanied by leaps in analytical power, the prime example being the Calculus.

What we need, then, is a criterion which is (i) <u>intuitively plausible</u>; (ii) <u>powerful</u>, in the sense that it enables us to decide for a large variety of arrangements which of a given pair is the better (ideally, it should enable us to order any two arrangements); (iii) <u>consistent</u>, in the sense that the preference relation established should contain no cycles; furthermore, (iv) in the special case where total costs are finite, it should <u>reduce to a simple total cost minimization criterion</u> and (v) in the special case of a covering of the plane by identical, autarchic regions, it should <u>reduce to the cost-density minimization criterion</u>, as noted above; finally, and not least, (vi) it should be <u>wieldy</u>, in the sense that one can decide which of two arrangements is better without excessive calculation.

Are there any criteria meeting these requirements? Yes, several, except that their wieldiness has yet to be tested. The best that we have been able to discover is the following:

<u>Arrangement L' is preferable to L" if there is a bounded region of the plane, H, such that, for any circular disc D containing H, the total cost incurred inside D under L' is less than the total cost incurred</u>

inside D under L".

This will be called the circles criterion.

Intuitively, the circles criterion proclaims one arrangement better than another if the first has lower costs in all sufficiently large regions of a certain regular character.[10] The criterion is powerful, though it does not succeed in ordering the set of possible arrangements completely. It can be shown to be consistent. It clearly boils down to a simple total cost comparison when arrangements with finite total costs are compared. In Lemma 3, below, it is proved that it reduces to a cost per unit area comparison in the case of arrangements partitioning the plane into identical, autarchic regions. Its wieldiness is demonstrated in the simple class of arrangements we have considered in Section 2, at least, and remains to be tested for more complex arrangements.[11]

4. APPLICATION OF THE CIRCLES CRITERION TO LATTICES

This section is intended to illustrate how the circles criterion might be used as a working principle, as well as to give some results which may lead to the strengthing of our theorem.

Let A stand for the dirichlet area of a lattice, and R for its dirichlet radius, as defined above, Section 2A.

---

10 Some regularity conditions must be imposed if the criterion is to have a reasonable degree of power. The class of circular discs accomplishes this, and is at the same time not too unwieldy in applications.

11 The same unboundedness problem has arisen in the case of economic development programs with an unlimited time horizon, and a similar criterion has been suggested for these cases (von Weizsäcker [15], and the entire issue [14]). The space problem is inherently more complex, however, because of multi-dimensionality, and the lack of anything corresponding to "time zero".

No point in a dirichlet region is at distance greater than R from the lattice point of that region, and no two points in the same region are at distance greater than 2R from each other. These simple observations are the keys to the following two lemmas.

Lemma 1. Let M be the number of lattice points enclosed in a circle of radius W ($>$R); then

(34) $$\pi(W + R)^2/A > M > \pi(W - R)^2/A.$$

Proof. The dirichlet regions of the M points are all enclosed in a circle of radius W + R concentric with the given circle. The total area covered by these regions is MA. Hence $MA < \pi(W + R)^2$, giving the first inequality. (The strictness of this and similar inequalities follows from the fact that hexagons cannot be fitted together to cover a circular region exactly.)

Now consider the dirichlet regions which overlap the circular disc of radius W - R concentric with the given circle. ("Overlap" is understood to include the case of complete enclosure of the region by the disc.) Each of these regions has its lattice point within the circle of radius W. Hence there are at most M of them, covering an area of at most MA. At the same time this area covers the inner circle, so that $MA > \pi(W - R)^2$, yielding the last inequality. QED

Lemma 2. Let N^e be the number of dirichlet regions enclosed in a circular disc of radius W ($>$2R); let N^o be the number of dirichlet regions overlapping that circular disc; then

(35) $$\pi(W + 2R)^2/A > N^o > N^e > \pi(W - 2R)^2/A .$$

Proof. Each of the N^o regions overlapping the disc of radius W must be enclosed by the concentric disc of radius W + 2R. Hence $N^o A < \pi (W + 2R)^2$, yielding the first inequality.

The middle inequality is obvious.

Now consider the dirichlet regions which overlap the circular disc of radius W - 2R concentric with the given disc. Each of these must be enclosed in the disc of radius W. Hence there are at most N^e of them, and their collective area covers the inner disc. Therefore $N^e A > \pi (W - 2R)^2$, yielding the last inequality. QED

We now use Lemma 2 to show that, for the simple class of arrangements used in our theorem, the circles criterion reduces to cost-density minimization.

Lemma 3. Let L_1 and L_2 be two lattices in which each plant is the exclusive supplier of its dirichlet region. Let A_1 and A_2 be the dirichlet areas of L_1 and L_2, respectively. Let C_1 and C_2 be the total costs (production plus transportation) incurred in a dirichlet region of L_1 and L_2, respectively. Then, if $C_1/A_1 < C_2/A_2$, L_1 is better than L_2 (by the circles criterion).

Proof. Let $C_1/A_1 = B_1$, $C_2/A_2 = B_2$. By assumption, $B_1 < B_2$.

We must find a bounded region, H, such that, on any circular disc enclosing H, total costs for L_1 are less than for L_2. For the region H we will choose any circular disc of radius W, where

$$(36) \qquad W > \frac{2R_2\sqrt{B_2} + 2R_1\sqrt{B_1}}{\sqrt{B_2} - \sqrt{B_1}},$$

R_1 and R_2 being the dirichlet radii of L_1 and L_2, respectively.

It suffices to show that total costs incurred are less under L_1 than under L_2 for any circular disc of radius W satisfying (36).

Let Z_1 and Z_2 be these total costs. Let N_1^o be the number of dirichlet regions of L_1 overlapping this disc, and let N_2^e be the number of dirichlet regions of L_2 enclosed in this disc. Then

(37) $$Z_2 \geq C_2 N_2^e \geq \pi(W - 2R_2)^2 C_2/A_2 = \pi(W - 2R_2)^2 B_2 \quad .$$

The first inequality arises from the fact that total cost over the whole disc is at least as large as over a portion of it, e.g. the portion consisting of all enclosed dirichlet regions. The second inequality comes from (35).

(38) $$Z_1 \leq C_1 N_1^o \leq \pi(W + 2R_1)^2 C_1/A_1 = \pi(W + 2R_1)^2 B_1 \cdot$$

The first inequality arises from the fact that the disc is covered by the totality of dirichlet regions overlapping it, the second inequality again from (35).

It remains only to complete this chain of inequalities. Now $(X - 2R_2) / (X + 2R_1)$ is an increasing function of X, so that, if we substitute from both sides of (36), we get

(39) $$\frac{W - 2R_2}{W + 2R_1} > \frac{2R_2\sqrt{B_2} + 2R_1\sqrt{B_1} - 2R_2 \ (\sqrt{B_2} - \sqrt{B_1})}{2R_2\sqrt{B_2} + 2R_1\sqrt{B_1} + 2R_1 \ (\sqrt{B_2} - \sqrt{B_1})} .$$

The right-hand side of (39) simplifies to $\sqrt{B_1} / \sqrt{B_2}$. From this follows at once

(40) $$\pi(W - 2R_2)^2 B_2 > \pi(W + 2R_1)^2 B_1 \quad .$$

(37), (38), and (40) imply that $Z_1 < Z_2$, which completes the proof.

Lemma 3 may be thought of as a justification for the cost-density minimization criterion, where it is applicable.

We now attempt a modest strengthening of our main theorem. The attempt almost succeeds, but not quite. Nonetheless, the partial results seem well worth presenting, both as an indication of what can be done, and of the surprising obstacles which can arise.

The modest strengthening contemplated is to broaden the class of possible arrangements as follows: (4a) is retained -- that is, we still confine ourselves to arrangements in which plants are arrayed in a lattice; but (4b) is dropped, and replaced merely by the assumption (4b') that production at all plants is the same. We no longer put any constraints on the pattern of transportation flows, save that demand requirements must be met, and of course, that total outflow from a plant equal its production level, which must be the same for all plants.

The original assumption (4b) -- that each plant be the exclusive supplier of its dirichlet region -- automatically fixed production at $A\rho$ (formula (1)) at all plants, and thus satisfied (4b'). This shows that the original class of arrangements is a proper subset of the new class defined by (4a) and (4b'). Thus, <u>if</u> we can derive the optimality of the honeycomb lattice from these new premises we will have strengthened our result.

Note that we can no longer use the cost-density criterion. Since transportation flow patterns are unrestricted, there need no longer be a "typical" region whose production-transportation pattern is repeated indefinitely, and thus cost-density will no longer be defined for all arrangements. We must fall back on the circles criterion.

Our first result states that, for a given lattice array of plants, the optimal common production level is $A\rho$, the output which would just supply a dirichlet region. This seems sort of obvious, but actually it is highly paradoxical, since nothing whatsoever is assumed about production costs (except that they are finite).

Note the difference in status of the formula $P = A\rho$ here, and in Section 2. There, it followed immediately from assumption (4b). Here it is derived as the result of an optimizing choice.

Lemma 4. Given a lattice array of plants, such that required deliveries in any dirichlet region is a constant, $A\rho$, and such that output at all plants, P, be equal, the optimal common production level is $P^* = A\rho$.

Proof. Let us first consider the choice $P = A\rho$. We have a large number of options as to the arrangement of transportation flows. One of them, clearly, is to have each plant be the exclusive supplier of its dirichlet region. This may not be the best arrangement, but it furnishes a convenient upper bound to the costs incurrable under the choice $P = A\rho$.

Let C be an upper bound for the total costs (production plus transportation) incurred in a dirichlet region under the exclusive supplier arrangement.

Now consider a circular disc of radius W, and let N^o be the number of dirichlet regions overlapping it, and Z the cost incurred on it under the choice $P = A\rho$. Then

$$Z \leq CN^o \leq \pi(W + 2R)^2 \, C/A, \tag{41}$$

where R as usual is the dirichlet radius of the lattice. The derivation of (41) is exactly the same as the derivation of (38).

The important fact about (41) is that the right hand term is a quadratic function of W. Let us rewrite it in the form

(42) $$Z \leq 0\ (W^2),$$

where $0\ (W^2)$ is a generic symbol indicating a function approaching a fixed positive ratio to W^2 as $W \rightarrow \infty$.

Next consider the choice $P < A\rho$. As in Lemma 1 let M stand for the number of lattice points enclosed in a circle.

It is clear for this case that any fair-sized region will incur a net deficit which has to be made up by imports. Let us find a lower bound for this deficit on a circle of radius V. Let D be total delivery requirements and S total supply forthcoming in this circle.

(43) $$D \geq A\rho N^e \geq \pi A\rho (V - 2R)^2/A\ ,$$

where N^e is the number of dirichlet regions enclosed in the circular disc. The justification for (43) is completely analogous to the justification for (37) above.

(44) $$S = PM \leq \pi P(V + R)^2/A.$$

The equality states that total production in the disc equals P times the number of plants in the disc. The inequality comes from (34).

Subtracting (44) from (43), we find, for the net deficit,

(45) $$D - S \geq \frac{\pi}{A}\ (A\rho - P)\ V^2 + \text{linear terms in } V = 0\ (V^2).$$

Thus the net deficit goes up as the square of the radius. This quantity must be imported through the rim of the circle. Consider the transportation cost incurred in a thin ring between two concentric circles, of radii V, and V + dV. It must be at least equal to (D - S) dV (greater, if traffic is flowing obliquely, or in two directions).

Let Z' be total costs incurred in a circle of radius W by a choice $P < A\rho$. Z' must at least equal the costs incurred by transportation alone, and so we have

$$(46) \qquad Z' \geq \int_0^W (D - S)\; dV \geq \int_0^W O(V^2)\; dV = O(W^3).$$

The first inequality arises from the integration of transportation costs over thin concentric shells, the second from (45). Comparing (42) and (46) we find that Z is dominated by a quadratic, while Z' dominates a cubic. Therefore Z' must eventually exceed Z' for all sufficiently large circles. This demonstrates the superiority of the choice $P = A\rho$ by the circles criterion.

The third possibility, $P > A\rho$, is handled in a manner analogous to the second. One finds a lower bound for the net surplus in a circular disc, using the other halves of the inequalities (34) and (35). Integration shows that in this case also, total costs dominate $O(W^3)$. Details are omitted. This completes the proof.

Lemma 4 did not assume that demand requirements were uniform, but merely that total requirements for each dirichlet region are equal.

As mentioned above, Lemma 4 is quite paradoxical. To take an extreme case, suppose that average cost of production equals 10 at all output levels except for $P = A\rho$, where, by some quirk, average cost equals 1000. If one had to choose equal outputs at all plants, the optimal choice would still be $P = A\rho$!

The paradox arises from the imposition of the fixed production level constraint. It would clearly be better in this case to vary production from plant to plant, avoiding the high production cost level, even if a little extra transportation cost is thereby incurred. Here is a case where a non-Christaller-Lösch type solution is better than one which does conform to the tradition. This does not refute the traditional solution, however, since another conforming arrangement may be better still.

We have merely assumed that production at all plants was equal, and concluded that it ought to be chosen just sufficient to supply a dirichlet region. It would seem fairly obvious that in this situation the optimal pattern of transportation flows would be for each plant to be the exclusive supplier of its dirichlet region, since each point would then be supplied from the closest available plant.

Obvious -- yet it is just at this point that we have gotten stuck. The trouble is that, of necessity, we must compare arrangements over (ever-increasing) finite portions of the plane, and, although the transportation costs generated in supplying the circular disc are minimized by the exclusive supplier arrangement, the transportation costs incurred on the disc need not be. We have not been able to overcome this difficulty, so that the optimality of the exclusive supplier arrangement is still unproved.

We do have some partial results which may prove useful to establishing the result. Let us enumerate all the lattice points in an infinite sequence: X_1, X_2, ---. Also enumerate all the dirichlet regions: D_1, D_2 --- in such a way that X_n is the lattice point of region D_n. A transportation flow may be (incompletely) characterized by the numbers $t(m,n)$, which is the quantity shipped from the plant at X_m to the region D_n.

Imagine the numbers $t(m,n)$ arrayed in a doubly-infinite matrix. The sum of the infinite series along each row, $\sum_n t(m,n)$ is the same, $A\rho$, since this is the output of each plant. Also, the sum of the series down each column is $A\rho$, since this is the requirement of each region. The exclusive supplier arrangement is exactly characterized by $t(m,n) = 0$ if $m \neq n$, $t(n,n) = A\rho$ for all n: an infinite scalar matrix.

A given flow is said to contain a <u>cycle</u> if the following is true: There are different numbers $m_1, m_2, ---, m_k$ $(k > 1)$ such that $t(m_1,m_2) > 0$, $t(m_2,m_3) > 0$, $---t(m_{k-1},m_k) > 0$, and finally $t(m_k,m_1) > 0$.

<u>Lemma 5</u>. A transportation flow, T, containing a cycle is non-optimal.

<u>Proof</u>. We construct another flow, T', which is superior, in the following way. Let b be the smallest of the numbers $t(m_i, m_{i+1})$, $(i = 1, ---,k)$, where m_{k+1} is identified with m_1. Let $t'(m_i, m_{i+1}) = t(m_i,m_{i+1}) - b$ $(i = 1, ---, k)$. Also let $t'(m_i, m_i) = t(m_i, m_i) + b$ $(i = 1, ---,k)$. For all other pairs m, n, one keeps $t'(m,n) = t(m,n)$. That is, one reduces the flows in the cycle all round by a positive constant, and increases the flow from each of these points to its own region by the same constant.

First, one confirms that this is an admissible flow by checking that $\sum_n t'(m,n,) = A\rho$ for all m, and $\sum_m t'(m,n) = A\rho$ for all n. This being established, one must now find a bounded region H such that in all circles containing H, the costs incurred under T' are less than under T.

For H we choose any circular disc containing $D_{m_1} \cup --- \cup D_{m_k}$. For any point in D_n, the distance to it from X_n is less than the distance to it from X_m, if $m \neq n$. Therefore, the transportation cost incurred in shipping a quantity from X_n to D_n is less than the cost incurred in shipping an equal quantity from X_m, if the distribution over destinations in D_n is the same. Thus, the saving from reducing the flow from X_{m_i} to $D_{m_{i+1}}$ by b exceeds the extra cost incurred in increasing the flow from $X_{m_{i+1}}$ to $D_{m_{i+1}}$ by b. All of these changes, and no others, are incurred in H, and all circles containing H, so T' is proved superior to T. QED

This result is partial, since there are lots of other flows for which $t(m,n) > 0$ for some $m \neq n$, and yet which contain no cycles. And even if these could all be proved inferior, this would still not establish the optimality of exclusive supply, since there might simply not be an optimal flow.

If the gap we have been discussing were closed, then the strengthened form of our theorem would follow at once. For, merely from the assumption (4b') of equality of outputs we would have proved successively (1) that these outputs were at level $A\rho$, and (2) that each plant should be the exclusive supplier of its dirichlet region. But this means we would have proved that (4b) was an optimality condition. From this point on the proof of Section 2 would be repeated verbatim.

5. THE POSSIBILITIES FOR FURTHER EXTENSIONS

Here we discuss, briefly and informally, other directions in which our theorem might be strengthened.

What about the possibility of dropping the restrictive assumption (3) of uniform delivery requirements, and going over to a general benefit function? It may be shown in this case that one gets a formation of "Thünen-rings" around each plant, density of deliveries dropping off with distance from the plant (possibly to zero near the corners of the dirichlet region, as pointed out by Mills and Lav [13]). These rings are cut off by the hexagonal sides of the region in a manner which is mathematically most unpleasant. In particular, it is no longer possible to separate the "size" from the "shape" problem as was done in Section 2B. One might venture the guess that any attempt to tred the same path as our proof would become hopelessly bogged down in complications. Our methods -- elementary calculus and trigonometry -- would then have to be replaced by deeper approaches.

These comments apply even more forcefully to the possibility of dropping the lattice assumption (4a). Here we must use the circles criterion or some equivalent. The dirichlet regions become convex polygons of any size or shape. If production at all plants does not just fill the requirements of their respective regions, the borderlines become hyporbolic segments. In short, the complications go up by several orders of magnitude. Again, deeper methods are required, probably based on generalized properties such as the convexity of transportation costs as a function of position, and symmetry.[12]

6. SUMMARY

We have proved the optimality of the honeycomb lattice in a fairly extensive class of cases, though the achievement is very modest in comparison with what remains to be done. We have also suggested, and applied to some extent, a new criterion of optimality to cope with the problem of infinite costs.

---

[12] Elementary examples of such approaches are given in my book [6].

Finally, we have tried to convey the serious mathematical nature, the gaps and loose ends, and the very limited knowledge possessed, of a problem which lies at the heart of location theory and spatial economics.[13]

---

[13] These same strictures apply to other branches of location theory as well -- for example, the theory of Thunen systems (see Faden[6]). They do not apply to "regional analysis", of which there has been a great efflorescence in recent years, and to which different standards of criticism apply. The difference in subject matter between these fields may be gleaned by comparing Isard 1956 [9] with Isard 1960 [10].

REFERENCES

[1] Cassels, J. W. S., <u>An Introduction to the Geometry of Numbers</u> (Berlin: Springer, 1959).

[2] Christaller, W., <u>Central Places in Southern Germany</u> (C. W. Baskin, trans.), (Englewood Cliffs, N.J.: Prentice-Hall, 1966).

[3] Cooper, L., "Solutions of Generalized Locational Equilibrium Models," <u>Journal of Regional Science</u>, vol. 7, #1, pp. 1-18, Summer, 1967.

[4] Coxeter, H. S. M., <u>Introduction to Geometry</u> (New York: Wiley, 1961)

[5] Dacey, M. F., "A Note on Some Number Properties of a Hexagonal Hierarchical Plane Lattice," <u>Journal of Regional Science</u>, vol. 5, #2, pp. 63-67, Winter, 1964.

[6] Faden, A. M., <u>Foundations of Spatial Economics</u> (Baltimore: Johns Hopkins University Press, forthcoming).

[7] Fejes-Tóth, L., <u>Lagerungen in der Ebene auf der Kugel und im Raum</u> (Berlin: Springer, 1953).

[8] Haggett, P., <u>Locational Analysis in Human Geography</u> (New York: St. Martin's Press, 1966).

[9] Isard, W., <u>Location and Space-Economy</u> (Cambridge, Mass.: Technology Press and Wiley, 1956).

[10] Isard, W., et al., <u>Methods of Regional Analysis</u> (Cambridge, Mass.: Technology Press and Wiley, 1960).

[11] Kuehn, A. A., and M. J. Hamburger, "A Heuristic Program for Locating Warehouses," <u>Management Science</u>, vol. 10, July, 1963.

(continued)

[12] Lösch, A. The Economics of Location (W. H. Woglom, trans.), (New Haven: Yale University Press, 1954).

[13] Mills, E. S., and M. R. Lav, "A Model of Market Areas with Free Entry," Journal of Political Economy, 72: 278-288, June, 1964.

[14] Review of Economic Studies, vol. 34, January, 1967.

[15] von Weizsäcker, C. C., "Existence of Optimal Programmes of Accumulation for an Infinite Time Horizon," Review of Economic Studies, 32: 85-104, April, 1965.