Math. Proc. Camb. Phil. Soc. (1995), **117**, 31
Printed in Great Britain

An elementary proof of the Birkhoff–Hopf theorem

BY SIMON P. EVESON
Department of Mathematics, University of York, Heslington, York, YO1 5DD

AND ROGER D. NUSSBAUM*
Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08903, U.S.A.

(*Received* 29 *June* 1993)

1. *Introduction*

In important work some thirty years ago, G. Birkhoff[**2**, **3**] and E. Hopf[**16**, **17**] showed that large classes of positive linear operators behave like contraction mappings with respect to certain 'almost' metrics. Hopf worked in a space of measurable functions and took as his 'almost' metric the oscillation $\omega(y/x)$ of functions y and x with $x(t) > 0$ almost everywhere, defined by

$$\omega(y/x) = \operatorname{ess.sup}_t \frac{y(t)}{x(t)} - \operatorname{ess.inf}_t \frac{y(t)}{x(t)}.$$

Birkhoff used what has been called ([**7**]) 'Hilbert's projective metric' or ([**9**]) the 'Cayley–Hilbert metric'. In each case, it proved possible to obtain sharp estimates for the contraction constant of a positive linear operator with respect to the 'almost' metric.

Subsequently, several authors generalized and sharpened the original results and established a close connection between the Birkhoff and Hopf theorems. A partial list of contributors includes F. L. Bauer[**1**], M. A. Ostrowski[**25**, **26**] and P. J. Bushell [**8**, **7**, **9**]. In addition, a number of mathematicians who were apparently unaware of most of the above-mentioned theorems obtained closely related results and interesting new propositions. We mention A. M. Krasnosel'skii, Je. E. Lifshits, Yu. V. Pokornyi, A. V. Sobolev, and refer the reader to [**19**], [**31**] and the book [**18**].

We shall prove here a generalization of the work of Birkhoff, Hopf, Bauer, Ostrowski, Bushell and others and refer to the cumulative result as the Birkhoff–Hopf Theorem: see Theorems 3·5 and 3·6 below and the formulae of Section 6.

Typically, when the Birkhoff–Hopf Theorem is applied to a positive (possibly non-compact) linear operator L, it implies that the L has a unique, normalized, positive eigenvector v with corresponding eigenvalue λ equal to the spectral radius of L and that there are explicitly computable constants M and c, with $c < 1$, such that

$$\left\| \frac{L^n x}{\|L^n x\|} - v \right\| \leqslant M c^n$$

* Partially supported by NSFDMS 91-05930.

for all positive vectors x. In addition, one may obtain explicitly computable formulae for the so-called spectral clearance $q(L)$ given by

$$q(L) = \sup\left\{\frac{|z|}{r(L)} : z \in \sigma(L)\backslash\{r(L)\}\right\},$$

where $\sigma(L)$ denotes the spectrum of L and $r(L)$ its spectral radius. Indeed, such estimates were one of the original motivations for the Birkhoff–Hopf theorem. We shall discuss these results more fully in a sequel to this paper [**13**].

In fact, the Birkhoff–Hopf theorem and associated ideas are important for a wide variety of problems. It plays a central role in proving so-called linear and nonlinear weak ergodic theorems of population biology (see [**10**, **14**, **15**, **23**]); it is a useful tool in problems concerned with rescaling matrices or non-negative integral kernels (so-called DAD theorems; see [**6**, **24**], [**22**, Section 4] and the references to the literature in [**22**]); it has proved crucial in some problems concerning ordinary differential equations, particularly the question of convergence in direction (see [**5**, **4**, **29**, **30**, **32**]); finally, these ideas play an important role in discussing convergence of $f^n(x)$ and $f^n(x)/\|f^n(x)\|$ when f belongs to an appropriate class of nonlinear operators and x is an element of a cone (see [**21**, **23**]).

Despite its usefulness, the Birkhoff–Hopf theorem is not as widely known as it should be, perhaps because of what A. M. Ostrowski [**25**, p. 91] has called a 'certain inaccessibility of Birkhoff's presentation'. As far as we know, we present here the first self-contained, elementary proof of the most general form of the theorem, treating a vector space V with a cone $C \subseteq V$, a vector space W with a cone $D \subseteq W$ and a linear map $L: V \to W$ with $L(C) \subseteq D$. Our basic observation is that it suffices to prove the theorem in the case when V and W are two-dimensional. Next we analyse two-dimensional cones and show that it suffices to prove the theorem when $V = W = \mathbb{R}^2$, $C = D = \{x \in \mathbb{R}^2 : x_1, x_2 \geqslant 0\}$ and $L = \left[\begin{smallmatrix} \alpha & 1 \\ 1 & \alpha \end{smallmatrix}\right]$, with $\alpha > 1$. Proving the theorem in this case is a simple calculus exercise which is carried out in Section 5. An amusing benefit of our proof is that, in contrast to all previous work, we need no assumption that our cones are Archimedean or almost Archimedean.

The approach here basically follows unpublished notes of R. D. Nussbaum which were written in 1986–87 and were one topic in a series of lectures at Emory University in the Spring of 1988. Independently, S. P. Eveson [**12**, **11**] found a closely related proof of the theorem for the almost Archimedean case. The present paper unites and refines these two approaches.

2. *Preliminary definitions and results*

Definition 2·1. If V is a real vector space and C is a subset of V, we shall call C a cone (with vertex at 0) if it satisfies the following three properties:

(1) C is convex;

(2) $tC \subseteq C$ whenever $t \geqslant 0$, where $tC = \{tx : x \in C\}$;

(3) $C \cap (-C) = \{0\}$.

If C satisfies properties (1) and (2), but not necessarily (3), we shall call C a wedge.

Remark 2·2. Note that we do not, in contrast to some of the literature, assume that V is a topological vector space in which C is closed. If V is a topological vector space and C is a cone in V, then the closure $\bar{C}$ of C need not be a cone (for example, let $V = \mathbb{R}^2$ and $C = \{(u,v): u > 0\} \cup \{(0,0)\}$). It is, however, easily verified that $\bar{C}$ is a wedge.

Definition 2·3. A cone C in a real vector space V induces a partial ordering on V by

$$x \leqslant_C y \quad \text{if and only if} \quad y - x \in C.$$

If there is no danger of confusion, we shall write $x \leqslant y$ instead of $x \leqslant_C y$.

If $x \in C$ and $y \in V$, we shall say that x dominates y if there exist real numbers α and β with

$$\alpha x \leqslant_C y \leqslant_C \beta x. \tag{1}$$

If in addition $x \neq 0$, we shall follow Bushell [7] and define

$$M(y/x; C) = \inf\{\beta \in \mathbb{R} : y \leqslant_C \beta x\},$$
$$m(y/x; C) = \sup\{\alpha \in \mathbb{R} : \alpha x \leqslant_C y\},$$
$$\omega(y/x; C) = M(y/x; C) - m(y/x; C),$$

which we shall abbreviate to $M(y/x)$, $m(y/x)$ and $\omega(y/x)$ when there is no danger of confusion. It follows from (1) and the hypothesis $x \neq 0$ that $m(y/x)$ and $M(y/x)$ are both finite. The quantity $\omega(y/x; C)$ is called the oscillation of y over x with respect to C; for convenience, we define $\omega(0/0; C) = 0$.

It is a straightforward exercise to show that if x dominates y then $m(y/x; C) \leqslant M(y/x; C)$, so $\omega(y/x; C) \geqslant 0$ and that if s and t are positive scalars then $M(sy/ty) = sM(y/x)/t$, $m(sy/ty) = sm(y/x)/t$ and hence that $\omega(sy/ty) = s\omega(y/x)/t$.

LEMMA 2·4. *Let V be a real vector space and let C be a cone in V. For any $x \in C\backslash\{0\}$, the set V_x of vectors in V dominated by x, that is*

$$V_x = \{y \in V : \alpha x \leqslant y \leqslant \beta x \text{ for some } \alpha, \beta \in \mathbb{R}\}$$

is a linear subspace of V. If we define

$$p_x(y) = \omega(y/x; C),$$

then p_x is a seminorm on V_x, which is to say that if y, $y_1, y_2 \in V_x$ and $\lambda \in \mathbb{R}$ then

$$p_x(y_1 + y_2) \leqslant p_x(y_1) + p_x(y_2); \quad p_x(\lambda y) = |\lambda| p_x(y).$$

It is also true that

$$\omega(\lambda y + \mu x/\nu x) = \frac{|\lambda|}{\nu}\omega(y/x)$$

for $y \in V_x$, λ, μ, $\nu \in \mathbb{R}$ and $\nu > 0$.

We leave the easy proofs of these facts to the reader. The subspace V_x is also the domain of the classical order norm $|\cdot|_x$ introduced by Krein and Rutman [**20**].

Remark 2·5. Of course, $\omega(y/x; C) = 0$ whenever $y = \lambda x$. However, we may have $\omega(y/x; C) = 0$ when y is not a multiple of x. To see this, consider $V = \mathbb{R}^2$ and $C = \{(u,v) \in \mathbb{R}^2 : u > 0\} \cup \{(0,0)\}$. The reader will easily verify that $M(y/x; C) = 0$ for all $x, y \in C\backslash\{0\}$.

If we consider $V = C(S)$, the Banach space of continuous functions on a compact Hausdorff space S, let C be the cone of non-negative functions in V and choose an everywhere positive function x, then we have

$$\omega(y/x; C) = \max_{s\in S} \frac{y(s)}{x(s)} - \min_{t\in S} \frac{y(t)}{x(t)}.$$

As a special case, where S is a discrete space of n points, we have $C(S) = \mathbb{R}^n$ and

$$\omega(y/x) = \max_{i=1}^{n} \frac{y_i}{x_i} - \min_{i=1}^{n} \frac{y_i}{x_i}.$$

A similar example, first studied by Hopf[**16**, **17**], arises in a space of real measurable functions, quotiented as usual by the equivalence relation of 'equal almost everywhere' and partially ordered by the cone of equivalence classes of almost everywhere non-negative functions. In this case, the supremum and infimum are replaced by their 'essential' counterparts.

Definition 2·6. Let C be a cone in a real vector space V. If $x, y \in C\backslash\{0\}$, we shall say that x is comparable to y in C if x dominates y and y dominates x. We shall write $x \sim_C y$ or $x \sim y$ to denote this, using the second form where there is no danger of confusion. It is easily verified that comparability is an equivalence relation on C; its equivalence classes are known as components.

If $x \sim_C y$, we have $0 < m(x/y) \leqslant M(x/y)$ and we define

$$d(x, y; C) = \log \frac{M(x/y; C)}{m(x/y; C)},$$

which we shall as usual write as $d(x, y)$ when there is no danger of confusion. It is clear from the definition that in any cone the origin is comparable only to itself; we make the convention that $d(0, 0) = 0$.

The function $d(\cdot, \cdot; C)$ is called the Hilbert projective metric or Cayley–Hilbert metric induced by the cone C.

LEMMA 2·7. *Let V be a real vector space and let C be a cone in V. If x, y and z are comparable elements of C then*

$$d(x, y) = d(y, x); \quad d(x, z) \leqslant d(x, y) + d(y, z)$$

and if λ, $\mu > 0$ then

$$d(\lambda x, \mu y) = d(x, y); \quad d(x, \lambda x) = 0.$$

The reader may easily supply the simple proofs of these facts, or refer to [**8**] for proofs of these and other related results in a slightly less general framework.

The example given in Remark 2·5 shows that if $d(x, y) = 0$ then it is not necessarily true that x is a scalar multiple of y.

Definition 2·8. Let V be a real vector space and $C \subseteq V$ be a cone. If $T \subseteq V$, we denote by $\mathrm{co}(T)$ the convex hull of T and if $\Gamma \subseteq C$ we define $\mathrm{diam}(\Gamma; C)$ to be the diameter of Γ with respect to the Hilbert's projective metric on C, so

$$\mathrm{diam}(\Gamma; C) = \sup\{d(x, y; C) : x, y \in \Gamma \text{ and } x \sim_C y\}.$$

The following proposition contains a number of fundamental properties of the projective metric.

PROPOSITION 2·9. *Let C be a cone in a real vector space V.*

(*a*) *If Q and S are subsets of C and*

$$\{sx\colon s \geqslant 0, x \in Q\} = \{sy\colon s \geqslant 0, y \in S\}$$

then $\operatorname{diam}(Q;C) = \operatorname{diam}(S;C)$.

(*b*) *If $T \subseteq C$ and all elements of $T\backslash\{0\}$ are comparable in C then* $\operatorname{diam}(\operatorname{co}(T);C) = \operatorname{diam}(T;C)$.

(*c*) *Let S be a convex set in C such that $M(v/u;C) > 0$ for all $u, v \in S\backslash\{0\}$ such that u dominates v in C. Assume that* $\operatorname{diam}(S;C) < \infty$. *Then any two elements of $S\backslash\{0\}$ are comparable in C.*

(*d*) *If V is a Hausdorff topological vector space and C is closed then for any set $Q \subseteq C$ such that all elements of $Q\backslash\{0\}$ are comparable in C, we have*

$$\operatorname{diam}(\bar{Q};C) = \operatorname{diam}(Q;C).$$

Proof. (*a*) This follows immediately from the fact that $d(sx, ty; C) = d(x, y; C)$ if $s, t > 0$ and x is comparable to y in C, and the convention that $d(0,0) = 0$.

(*b*) It suffices to prove that $R = \operatorname{diam}(T;C) \geqslant \operatorname{diam}(\operatorname{co}(T);C)$, since the opposite inequality is obvious. Similarly, we may assume that $R < \infty$. Select $R_1 > R$.

If $v \in T\backslash\{0\}$ and $u \in \operatorname{co}(T)\backslash\{0\}$, a little thought shows that

$$u = t \sum_{k=1}^{n} \lambda_k u_k,$$

where $0 < t \leqslant 1$, $\sum_{k=1}^{n} \lambda_k = 1$ and for $1 \leqslant k \leqslant n$, $\lambda_k > 0$ and $u_k \in T\backslash\{0\}$.

By assumption, all non-zero elements of T are comparable, so there exist α_k, $\beta_k > 0$ such that $\alpha_k v \leqslant u_k \leqslant \beta_k v$ and $\log(\beta_k/\alpha_k) < R_1$. These inequalities yield

$$t\left(\sum_{k=1}^{n} \lambda_k \alpha_k\right) v \leqslant u \leqslant t\left(\sum_{k=1}^{n} \lambda_k \beta_k\right) v;$$

so

$$\begin{aligned} d(u,v) &\leqslant \log\left(\frac{\sum_{k=1}^{n} \lambda_k \beta_k}{\sum_{k=1}^{n} \lambda_k \alpha_k}\right) \\ &\leqslant \log\left(\frac{\sum_{k=1}^{n} \lambda_k \alpha_k e^{R_1}}{\sum_{k=1}^{n} \lambda_k \alpha_k}\right). \\ &= R_1. \end{aligned}$$

Since R_1 was any number greater than R, we conclude that $d(u,v) \leqslant R$ for all $u \in \operatorname{co}(T)\backslash\{0\}$, $v \in T\backslash\{0\}$.

Now suppose $u, v \in \text{co}(T)\backslash\{0\}$. We may write

$$v = s \sum_{k=1}^{m} \gamma_k v_k,$$

where $0 < s \leqslant 1$, $\sum_{k=1}^{m} \gamma_k = 1$ and for $1 \leqslant k \leqslant m$, $0 < \gamma_k \leqslant 1$ and $v_k \in T\backslash\{0\}$. Select $R_1 > R$. Since by the first part $d(u, v_k) \leqslant R$, there exist $\alpha'_k, \beta'_k > 0$ such that for $1 \leqslant k \leqslant m$, $\alpha'_k u \leqslant v_k \leqslant \beta'_k u$ and $\log(\alpha'_k/\beta'_k) < R_1$.

Just as before, this gives

$$s\left(\sum_{k=1}^{m} \gamma_k \alpha'_k\right) u \leqslant s \sum_{k=1}^{m} \gamma_k v_k = v \leqslant s e^{R_1}\left(\sum_{k=1}^{m} \gamma_k \alpha'_k\right) u,$$

which implies that $d(u, v) < R_1$. Since R_1 was an arbitrary number greater than R, we conclude that $\text{diam}(\text{co}(T); C) \leqslant R$.

(*c*) Assume for a contradiction that there exist non-zero elements u and v of S which are incomparable in C. We may assume without loss of generality that v does not dominate u in C, and consider two subcases: (1) that u does not dominate v in C and (2) that u does dominate v in C.

For $0 < \epsilon < 1$, define $u_\epsilon = u + \epsilon v$ and $v_\epsilon = v + \epsilon u$. Suppose first that we are in subcase (1). We claim that $M(v_\epsilon/u_\epsilon) = \epsilon^{-1}$ and $m(v_\epsilon/u_\epsilon) = \epsilon$. Certainly we have

$$v_\epsilon = v + \epsilon u \leqslant \epsilon^{-1}(u + \epsilon v) = \epsilon^{-1} u_\epsilon,$$

so $M(v_\epsilon/u_\epsilon) \leqslant \epsilon^{-1}$. On the other hand, if $M(v_\epsilon/u_\epsilon) < \gamma < \epsilon^{-1}$, we find that

$$(1 - \gamma\epsilon) v \leqslant (\gamma - \epsilon) u$$

and, since $1 - \gamma\epsilon > 0$, this implies that $\epsilon < \gamma$ and u dominates v, a contradiction. The argument that $m(v_\epsilon/u_\epsilon) = \epsilon$ is similar and is left to the reader.

Since S is convex, we have that $(1+\epsilon)^{-1} u_\epsilon \in S\backslash\{0\}$ and $(1+\epsilon)^{-1} v_\epsilon \in S\backslash\{0\}$. Finally,

$$d(u_\epsilon, v_\epsilon) = d((1+\epsilon)^{-1} u_\epsilon, (1+\epsilon)^{-1} v_\epsilon) = \log(\epsilon^{-2}).$$

Letting $\epsilon \downarrow 0$, we contradict the assumption that $\text{diam}(S; C) < \infty$.

We may therefore assume we are in subcase (2). Let $\beta = M(v/u) > 0$. Similar arguments to those above show that $M(v_\epsilon/u) = \beta + \epsilon$ and $m(v_\epsilon/u) = \epsilon$, so

$$d(v_\epsilon, u) = d((1+\epsilon)^{-1} v_\epsilon, u) = \log \frac{\beta + \epsilon}{\epsilon}.$$

Since $\beta > 0$, we again contradict the assumption that $\text{diam}(S; C) < \infty$.

(*d*) It suffices to show that $\text{diam}(Q; C) \geqslant \text{diam}(\bar{Q}; C)$, the opposite inequality being obvious. We may assume $\text{diam}\,(Q; C) < \infty$. If $u, v \in \bar{Q}\backslash\{0\}$ and u and v are comparable in C, let $\beta = M(v/u)$, $\alpha = m(v/u)$ and select $\alpha' > \alpha$ and $\beta' < \beta$. By definition of α and β, we have

$$\beta' u - v \notin C, \quad v - \alpha' u \notin C.$$

Since C is closed, there exist open neighbourhoods $\mathscr{U}$ of u and $\mathscr{V}$ of v with $0 \notin \mathscr{U} \cup \mathscr{V}$ and

$$\beta' u_1 - v_1 \notin C, \quad v_1 - \alpha' u_1 \notin C, \tag{2}$$

for all $u_1 \in \mathscr{U}$ and $v_1 \in \mathscr{V}$. By definition, there exists $u_1 \in Q \cap \mathscr{U}$ and $v_1 \in Q \cap \mathscr{V}$; by hypothesis u_1 is comparable to v_1 and by (2), we have

$$M(v_1/u_1) \geqslant \beta', \quad m(v_1/u_1) \leqslant \alpha',$$

which by letting $\beta' \to \beta$ and $\alpha' \to \alpha$ implies that $\operatorname{diam}(Q;C) \geqslant d(u,v)$. Since u and v were arbitrary comparable elements of $\bar{Q}\backslash\{0\}$, the result follows. ∎

Remark 2·10. Part (*c*) of Proposition 2·9 is false if one only assumes that S is convex and $\operatorname{diam}(S;C) < \infty$. To see this, take

$$C = \{x \in \mathbb{R}^2 : x_1 > 0 \text{ or } x_1 = 0 \text{ and } x_2 \leqslant 0\}.$$

If we define $S = C$, the reader will easily verify that $\operatorname{diam}(S;C) = 0$ but that $(0, -1)$ and $(1, 0)$ are incomparable elements of S.

Remark 2·11. In Definition 3·9 below, we define $V(x,y)$ to be the subspace spanned by x and y (so $V(x,y)$ has a natural topology) and define $C(x,y) = C \cap V(x,y)$. With this notation, one may easily verify that the condition in part (*c*) of Proposition 2·9 is satisfied if $\overline{C(u,v)}$ is a cone whenever $u, v \in S\backslash\{0\}$ and u dominates v in C.

3. *Positive linear operators*

If C and D are cones in real vector spaces V and W respectively and L is a linear map from V to W with $L(C) \subseteq D$, we wish to compare $\omega(y/x;C)$ with $\omega(Ly/Lx;D)$ and $d(x,y;C)$ with $d(Lx,Ly;D)$. Our first lemma is trivial but will play an important role.

LEMMA 3·1. *Suppose that C is a cone in a real vector space V, that D is a cone in a real vector space W and that $L: V \to W$ is a linear map with $L(C) \subseteq D$. If $x \in C$, $y \in V$ and x dominates y then Lx dominates Ly in W and*

$$\omega(Ly/Lx;D) \leqslant \omega(y/x;C). \tag{3}$$

If x, $y \in C$ and $x \sim_C y$ then $Lx \sim_D Ly$ and

$$d(Lx, Ly) \leqslant d(x,y). \tag{4}$$

If L is a bijection and $L(C) = D$ then equality holds in both (3) *and* (4).

Proof. If x dominates y in V and $\alpha x \leqslant_C y \leqslant_C \beta x$, then $\beta x - y \in C$ and $y - \alpha x \in C$, so $L(\beta x - y) \in D$ and $L(y - \alpha x) \in D$. In terms of the order relation, we have

$$\alpha Lx \leqslant_D Ly \leqslant_D \beta Lx, \tag{5}$$

showing that Lx dominates Ly in D. Now, if $Lx = 0$ then $Ly = 0$ so $\omega(Ly/Lx) = 0$ by definition and $\omega(Ly/Lx) \leqslant \omega(y/x)$ since the oscillation is always non-negative.

If, on the other hand, $Lx \neq 0$ then it follows from (5) that $M(Ly/Lx) \leqslant M(y/x)$ and $m(Ly/Lx) \geqslant m(y/x)$, and hence that (3) is true.

A similar approach establishes the corresponding results for d.

If L is a bijection between V and W with $L(C) = D$ then we can apply the results above to the map $L^{-1}: W \to V$, since $L^{-1}(D) \subseteq C$, and conclude that, for $x \in C$ and $y \in V$ with x dominating y,

$$\omega(L^{-1}(Ly)/L^{-1}(Lx)) \leqslant \omega(Ly/Lx)$$

and for $x, y \in C$ with x comparable to y,

$$d(L^{-1}(Lx), L^{-1}(Ly)) \leqslant d(Lx, Ly)$$

so equality holds in (3) and (4). ∎

Remark 3·2. If $V = W = \mathbb{R}^n$, $C = D = \{x \in \mathbb{R}^n : x_i \geqslant 0 \text{ for all } i\}$, and $L: \mathbb{R}^n \to \mathbb{R}^n$ is given by a positive diagonal matrix or a permutation matrix, then Lemma 3·1 implies that for any $x, y \in C$ with $x \sim y$,

$$d(Lx, Ly) = d(x, y)$$

and for any $x \in C, y \in V$ with x dominating y,

$$\omega(Ly/Lx) = \omega(y/x).$$

Definition 3·3. Let C be a cone in a real vector space V and D be a cone in a real vector space W. Define non-negative real numbers $N(L; C, D)$, $k(L; C, D)$, $\Delta(L; C, D)$ and $\chi(L; C, D)$ by

$$N(L; C, D) = \inf\{\mu \geqslant 0 : \omega(Ly/Lx; D) \leqslant \mu\omega(y/x; C) \text{ for all } x \in C, y \in V \text{ such that } x \text{ dominates } y\},$$

$$k(L; C, D) = \inf\{\lambda \geqslant 0 : d(Lx, Ly; D) \leqslant \lambda d(x, y; C) \text{ for all } x, y \in C \text{ such that } x \sim_C y\},$$

$$\Delta(L; C, D) = \sup\{d(Lx, Ly; D) : x, y \in C \text{ and } Lx \sim_D Ly\},$$

$$\chi(L; C, D)^2 = \sup\left\{\frac{M(Ly/Lx)}{m(Ly/Lx)} : x, y \in C, Lx \sim_D Ly\right\},$$

where in the definition of χ we adopt the convention that $M(Ly/Lx)/m(Ly/Lx) = 1$ if both $M(Ly/Lx)$ and $m(Ly/Lx)$ are zero.

$N(L; C, D)$ is called the Hopf oscillation ratio, $k(L; C, D)$ is called the Birkhoff contraction ratio and $\Delta(L; C, D)$ is called the projective diameter.

We shall as usual abbreviate these to $N(L)$, $k(L)$, $\Delta(L)$ and $\chi(L)$ if there is no danger of confusion. Lemma 3·4 implies that $k(L) \leqslant 1$ and $N(L) \leqslant 1$ and it is clear that $\Delta(L) \geqslant 0$ and $\chi(L) \geqslant 1$. We may have $\Delta(L) = \infty$ and $\chi(L) = \infty$, and if we let $\log \infty = \infty$, then we have

$$\Delta(L) = 2 \log \chi(L).$$

In much of the literature, $N(L; C, D)$ is defined only in terms of comparable elements of the cone. Our definition is equivalent, as the following lemma shows.

LEMMA 3·4. *In the notation of Definition* 3·3,

$$N(L; C, D) = \inf\{\mu \geqslant 0 : \omega(Ly/Lx; D) \leqslant \mu\omega(y/x; C) \text{ for all } x, y \in C \text{ such that } x \sim_C y\}.$$

Proof. This formula is identical to the definition of $N(L; C, D)$ except that y is restricted to being comparable to x instead of being allowed to be dominated by x. To prove it, we shall show that if $x \in C$ and $y \in V$ with x dominating y then there exists $z \in C$ comparable to x with $\omega(y/x) = \omega(z/x)$ and $\omega(Ly/Lx) = \omega(Lz/Lx)$.

To find such a z, let $\alpha x \leqslant y \leqslant \beta x$ and let $z = y + (1-\alpha)x$. We have $z \in C$ and $z \sim_C x$ since

$$x \leqslant z \leqslant (\beta - \alpha + 1)x$$

and by Lemma 2·4, $\omega(z/x) = \omega(y/x)$ and $\omega(Lz/Lx) = \omega(Ly/Lx)$. ∎

We shall generally use this formulation of $N(L; C, D)$ in preference to our original definition, but it is useful in applications not to restrict ourselves to non-negative vectors.

These definitions leave us in a position to state our main theorem.

THEOREM 3·5. *Let C be a cone in a real vector space V, D be a cone in a real vector space W and L be a linear map from V to W with $L(C) \subseteq D$. Suppose that $\Delta(L) < \infty$. Then*

$$k(L; C, D) = N(L; C, D) = \tanh \tfrac{1}{4}\Delta(L; C, D) = \frac{\chi(L; C, D) - 1}{\chi(L; C, D) + 1}.$$

THEOREM 3·6. *Let C be a cone in a real vector space V, D be a cone in a real vector space W and L be a linear map from V to W with $L(C) \subseteq D$. Suppose that $\Delta(L) = \infty$. Then*

$$k(L; C, D) = N(L; C, D) = 1.$$

We shall see later that it is easy to deduce Theorem 3·6 from Theorem 3·5.

These theorems have a long history. In the literature it has usually been assumed that $C = D$, but for various applications to nonlinear problems it is useful to have $C \neq D$ (see [**22**, section 4], [**24**] and [**6**]). Furthermore, as we shall see, allowing different cones permits a flexibility in our arguments which will actually simplify the proof.

It has also usually been assumed in the literature that V and W are topological vector spaces and that C and D are closed, or at least satisfy some sort of 'Archimedean' property. Our arguments will show that these assumptions are unnecessary.

If $V = W$ is a Banach space and $C = D$ is a closed cone in V, G. Birkhoff [**2**, **3**] showed that

$$k(L; C, C) = \tanh \tfrac{1}{4}\Delta(L; C, C).$$

E. Hopf proved (basically for the case where $V = W$ is a space of measurable functions like L^∞, $C = D$ is the cone of almost everywhere non-negative functions and L is an integral operator) that

$$N(L; C, C) \leqslant \tanh \tfrac{1}{4}\Delta(L; C, C),$$

though he never explicitly defined $\Delta(L; C, C)$. A different proof of a generalized version of Hopf's result was given by F. L. Bauer [**1**]. A. M. Ostrowski [**25**] and P. J. Bushell [**8**] showed that

$$N(L; C, C) = k(L; C, C).$$

Apparently unaware of some of the literature (notably Birkhoff's papers [**2**] and [**3**]), Zabreiko, Krasnosel'skii and Pokornyi [**31**] and Krasnosel'skii and Sobolev [**19**] have obtained closely related results.

The first step in proving Theorem 3·5, and the heart of our approach to the problem, is to show that it suffices to prove it when V and W have dimension less than or equal to 2. First, we need to give some definitions and recall some elementary results.

Definition 3.7. If C is a cone in a real vector space V, we shall call C finite-dimensional if there exists a finite-dimensional subspace E of V with $C \subseteq E$. In this case, we define $\dim(C)$, the dimension of C, to be the smallest dimension of a finite-dimensional linear subspace E of V with $C \subseteq E$.

Recall (see [**27**, chapter 1]) that if E is a finite-dimensional, real vector space, then there is a unique topology on E which makes E a Hausdorff topological vector space. It follows that if C is a finite-dimensional cone in a real vector space V, we may define $\bar{C}$, the closure of C, even though V is not assumed to be a topological vector space: we let E be any finite-dimensional subspace of V with $C \subseteq E$, take the unique topology on E which makes it into a Hausdorff topological vector space, and define $\bar{C}$ to be the closure of C in this topology. It is not hard to see that $\bar{C}$ is independent of the particular subspace chosen.

By virtue of this remark, topology will play some role in the proof of Theorem 3·5. We shall need another well-known result of point set topology.

LEMMA 3·8. *Let C be a convex set in a topological vector space E. If $x \in \mathring{C}$ and $y \in \bar{C}$ then $(1-t)x + ty \in \mathring{C}$ for $0 \leqslant t < 1$.*

Proof. See [**28**, chapter II, section 1·1]. ∎

By using Lemma 3·8, one may easily see that if C is a cone with non-empty interior in a topological vector space E and if $x \in \mathring{C}$ and $y \in \bar{C}$ then $\alpha x + \beta y \in \mathring{C}$ for all $\alpha > 0$ and $\beta \geqslant 0$. We shall use this fact later.

Definition 3·9. Let V be a real vector space and C a cone in E. If $x, y \in V$, then let

$$V(x,y) = \{\alpha x + \beta y : \alpha, \beta \in \mathbb{R}\},$$
$$C(x,y) = C \cap V(x,y).$$

Remark 3·10. In this definition, we clearly have $\dim(C) \leqslant 2$. Moreover, if $u, v \in C(x,y) =: K$, we have

$$u \leqslant_K v \quad \text{if and only if} \quad u \leqslant_C v.$$

It follows that if $u \in K$ and $v \in V$ then u dominates v with respect to K if and only if u dominates v with respect to C and that in this case

$$M(v/u; K) = M(v/u; C); \quad m(v/u; K) = m(v/u; C)$$

and hence that $\omega(v/u; K) = \omega(v/u; C)$. Similarly, one sees that for $u, v \in K$, u is comparable to v in K if and only if u is comparable to v in C and that

$$d(u,v;C) = d(u,v;K).$$

By using these elementary remarks, we can reduce the proof of Theorem 3·5 to the proof of a much simpler result.

LEMMA 3·11. *In order to prove Theorem* 3·5, *it suffices to prove it when* $\dim(C) = \dim(D) = 2$, $\dim(V) = \dim(W) = 2$ *and L is injective.*

Proof. Let V, W, C, D and L be as in Theorem 3·5.

If $\dim(C) \leqslant 1$ or $\dim(D) \leqslant 1$, one may easily verify that $k(L;C,D) = N(L;C,D) = \Delta(L;C,D) = 0$, so Theorem 3·5 is trivially true.

If $\dim(C) = \dim(D) = 2$ and $\dim(V) = \dim(W) = 2$ but L is not injective, then the range of L is of dimension zero or one; in either case, one may easily verify that $k(L;C,D) = N(L;C,D) = \Delta(L;C,D) = 0$, so Theorem 3·5 is again trivially true.

We now assume that Theorem 3·5 is true whenever the spaces and cones involved are two-dimensional and the map is injective. By virtue of the remarks above, this implies that it is true whenever the two spaces are two-dimensional.

We now abbreviate $k(L;C,D)$, $N(L;C,D)$ and $\Delta(L;C,D)$ to $k(L)$, $N(L)$ and $\Delta(L)$. For $x, y \in C$ we have (using the notation in Definition 3·9)

$$L\colon V(x,y) \to W(Lx,Ly); \quad L(C(x,y)) \subseteq D(Lx,Ly).$$

We may thus define functions k, N and Δ on $C \times C$ by

$$\begin{aligned} k(x,y) &= k(L;C(x,y),D(Lx,Ly)),\\ N(x,y) &= N(L;C(x,y),D(Lx,Ly)),\\ \Delta(x,y) &= \Delta(L;C(x,y),D(Lx,Ly)). \end{aligned}$$

Since we have assumed Theorem 3·5 to be true for spaces of dimension 2 or less, we have

$$k(x,y) = N(x,y) = \tanh \tfrac{1}{4}\Delta(x,y). \tag{6}$$

On the other hand, using the identities in Remark 3·10, we have

$k(x,y) = \inf\{\lambda \geqslant 0\colon d(Lu,Lv;D) \leqslant \lambda d(u,v;C)$ for all $u, v \in C(x,y)$ with $u \sim_C y\}$,

$N(x,y) = \inf\{\mu \geqslant 0\colon \omega(Lv/Lu;D) \leqslant \mu\omega(v/u;C)$ for all $u, v \in C(x,y)$ such that u dominates v in $C\}$,

$\Delta(x,y) = \sup\{d(Lu,Lv;D)\colon u, v \in C(x,y)$ and $Lu \sim_D Lv\}$.

It follows easily that

$$\begin{aligned} k(L) &= \sup\{k(x,y)\colon x, y \in C\},\\ N(L) &= \sup\{N(x,y)\colon x, y \in C\},\\ \Delta(L) &= \sup\{\Delta(x,y)\colon x, y \in C\}. \end{aligned}$$

By using these equations and (6), we see that

$$\begin{aligned} k(L) &= \sup\{k(x,y)\colon x, y \in C\}\\ &= \sup\{N(x,y)\colon x, y \in C\}\\ &= N(L)\\ &= \sup\{\tanh \tfrac{1}{4}\Delta(x,y)\colon x, y \in C\}\\ &= \tanh(\sup\{\tfrac{1}{4}\Delta(x,y)\colon x, y \in C\})\\ &= \tanh \tfrac{1}{4}\Delta(L). \quad \blacksquare \end{aligned}$$

In addition to Lemma 3·11, another tool to simplify the proof of Theorem 3·5 will be the replacement of the linear map L with a map SLT, where S and T are suitably chosen linear bijections.

LEMMA 3·12. *Let C be a cone in a real vector space V, D be a cone in a real vector space W and L a linear map from V to W with $L(C) \subseteq D$ and $\Delta(L; C, D) < \infty$. Let V_1 and W_1 be real vector spaces and $S: V_1 \to V$ and $T: W \to W_1$ be linear bijections. Define cones C_1 and D_1 by $C_1 = S^{-1}(C)$ and $D_1 = T(D)$ and let $L_1 = TLS$. Then*

$$k(L; C, D) = k(L_1; C_1, D_1),$$
$$N(L; C, D) = N(L_1; C_1, D_1),$$
$$\Delta(L; C, D) = \Delta(L_1; C_1, D_1).$$

In particular, to prove the identity in Theorem 3·5 it is sufficient to prove it for the map L_1 and the cones C_1 and D_1.

Proof. Lemma 3·1 implies that $\omega(Sv/Su; C) = \omega(v/u; C_1)$ and $d(u, v; C_1) = d(Su, Sv; C)$ for all appropriate $u, v \in C_1$. Similarly, we have $d(T\xi, T\eta; D_1) = d(\xi, \eta; D)$ and $\omega(T\eta/T\xi; D_1) = \omega(\eta/\xi; D)$ for all appropriate $\xi, \eta \in D$. The identities claimed all follow immediately from this. ▌

4. *Classification of two-dimensional cones*

In this Section, we shall show that the closure of a two-dimensional cone in the plane is either a half-plane or may be identified by means of a linear isomorphism with the positive quadrant. These representations give a simple formula for the projective metric in a two-dimensional cone. These facts are almost self-evident, but their proof does not appear to be as simple as one might expect.

The topology used on the plane throughout this section will be the usual Euclidean topology.

THEOREM 4·1. *Let C be a two-dimensional cone in $\mathbb{R}^2$. Then exactly one of the following two alternatives is true.*

(1) *$\bar{C}$ is a cone and there exist linearly independent vectors u and v such that*

$$\bar{C} = \{\lambda u + \mu v : \lambda, \mu \geqslant 0\},$$
$$\mathring{C} = \{\lambda u + \mu v : \lambda, \mu > 0\}.$$

(2) *$\bar{C}$ is not a cone and there exist linearly independent vectors u and v such that*

$$\bar{C} = \{\lambda u + \mu v : \lambda \in \mathbb{R}, \mu \geqslant 0\},$$
$$\mathring{C} = \{\lambda u + \mu v : \lambda \in \mathbb{R}, \mu > 0\}.$$

Proof. Since C is two-dimensional, it contains two linearly independent vectors, a and b. It follows that

$$\{\lambda a + \mu b : \lambda, \mu > 0\}$$

is an open set contained in C, so $\lambda a + \mu b \in \mathring{C}$ whenever $\lambda, \mu > 0$. In particular, there is a disc centred at $(a+b)$ and contained in C, so there is a disc centred at $-(a+b)$ entirely disjoint from C, and hence $-(a+b) \notin \bar{C}$.

We define $x = 2a+b\in\mathring{C}$ and $y = -(a+b)\notin\bar{C}$, and note that x and y are linearly independent. We now consider two cases: that $\bar{C}$ is a cone, and that it is not a cone.

If $\bar{C}$ is a cone, define τ by

$$\tau = \sup\{t\in(0,1): (1-t)y+tx\notin\bar{C}\}.$$

Our selection of x and y ensures that $0<\tau<1$, that $u = (1-\tau)x+\tau y\in\partial C$, and that (because x and y are linearly independent) $u \neq 0$. Since we are assuming that $\bar{C}$ is a cone, $-u\notin\bar{C}$. We now define σ by

$$\sigma = \sup\{s\in(0,1): (1-s)(-u)+sx\notin\bar{C}\}.$$

As before, we find that $0<\sigma<1$, that $v = (1-\sigma)(-u)+\sigma x\in\partial C$ and, since x and y are linearly independent, that u and v are also linearly independent.

Since $\bar{C}$ is a wedge, we have

$$\{\lambda u+\mu v: \lambda,\mu\geqslant 0\}\subseteq\bar{C}.$$

We easily derive from Lemma 3·8 that $\mathring{\bar{C}} = \mathring{C}$, so

$$\mathring{\bar{C}} = \mathring{C}\supseteq\{\lambda u+\mu v: \lambda,\mu>0\}.$$

To complete the proof, it suffices to show that

$$\bar{C} = \{\lambda u+\mu v: \lambda,\mu\geqslant 0\}.$$

If not, then there exists a point $z = \lambda_0 u+\mu_0 v\in\bar{C}$ with either $\lambda_0<0$ or $\mu_0<0$. We may assume without loss of generality that $\lambda_0<0$ and note that then $\mu_0>0$, since otherwise we would have $z\neq 0$, $z\in\bar{C}$ and $-z\in\bar{C}$. We now have, by Lemma 3·8, that

$$\tfrac{1}{2}(\lambda_0 u+\mu_0 v)+\tfrac{1}{2}(|\lambda_0|u+|\lambda_0|v) = \mu_0 v\in\mathring{C},$$

contradicting the fact that $v\in\partial C$.

We now consider the alternative case, that $\bar{C}$ is not a cone. Since $\bar{C}$ is a wedge, there exists $u\neq 0$ with $u, -u\in\bar{C}$. It follows that $u\in\partial C$, since otherwise we would have $u\in\mathring{C} = \mathring{\bar{C}}$ and $-u\in\bar{C}$, from which it would follow by Lemma 3·8 that

$$0 = \tfrac{1}{2}(u+-u)\in\mathring{\bar{C}} = \mathring{C}.$$

Similarly, $-u\in\partial C$.

Now, select $v\in\mathring{C}$, and note that v and u are linearly independent, since otherwise v would be a scalar multiple of u, and hence an element of ∂C. Because $\bar{C}$ is a wedge containing u, $-u$ and v,

$$\bar{C}\supseteq\{\lambda u+\mu v: \lambda\in\mathbb{R},\mu\geqslant 0\}.$$

To complete the proof, we must show that $\bar{C}$ contains no element of the form $\lambda_0 u+\mu_0 v$, where $\mu_0<0$. If $\bar{C}$ does contain such an element, then $\lambda_0 u+|\mu_0|v\in\mathring{C}$, so by Lemma 3·8 we have

$$\lambda_0 u = \tfrac{1}{2}(\lambda_0 u+\mu_0 v)+\tfrac{1}{2}(\lambda_0 u+|\mu_0|v)\in\mathring{C}$$

contradicting the fact that $u\in\partial C$. ∎

Corollary 4·2. (*Using the same notation and hypotheses as Theorem* 4·1.)

If $\bar{C}$ is a cone and x and y are elements of $\mathring{C}$, so

$$x = \lambda_1 u + \lambda_2 v, \quad y = \mu_1 u + \mu_2 v,$$

with $\lambda_1, \lambda_2, \mu_1, \mu_2 > 0$, then x is comparable to y and

$$M(x/y) = \max\left\{\frac{\lambda_1}{\mu_1}, \frac{\lambda_2}{\mu_2}\right\},$$

$$m(x/y) = \min\left\{\frac{\lambda_1}{\mu_1}, \frac{\lambda_2}{\mu_2}\right\},$$

$$\omega(x/y) = \left|\frac{\lambda_1}{\mu_1} - \frac{\lambda_2}{\mu_2}\right|,$$

$$d(x,y) = \left|\log\frac{\lambda_1 \mu_2}{\lambda_2 \mu_1}\right|.$$

If, on the other hand, $\bar{C}$ is not a cone then if $x \in C$, $y \in V$ and x dominates y then $\omega(y/x; C) = 0$, and if x is comparable to y in C, then $d(x,y;C) = 0$.

Proof. The proofs of the results which apply when $\bar{C}$ is a cone are left as a simple exercise for the reader. The proof of the remainder of the corollary follows.

Suppose x dominates y in the order relation induced by C. By Theorem 4·1, we have

$$x = \lambda_1 u + \lambda_2 v,$$

$$y = \mu_1 u + \mu_2 v.$$

Now, suppose

$$\alpha x \leqslant_C y \leqslant_C \beta x.$$

This is equivalent to

$$(\mu_1 - \alpha\lambda_1) u + (\mu_2 - \alpha\lambda_2) v \in C,$$

$$(\beta\lambda_1 - \mu_1) u + (\beta\lambda_2 - \mu_2) v \in C.$$

Thus, by the representation of $\bar{C}$, $\alpha\lambda_2 \leqslant \mu_2 \leqslant \beta\lambda_2$. Thus, if $\lambda_2 = 0$ then $\mu_2 = 0$ and both x and y are multiples of u, so $\omega(y/x) = 0$ and if they are comparable then $d(x,y) = 0$.

We shall now assume that $\lambda_2 > 0$, and show that $m(y/x) = M(y/x) = \mu_2/\lambda_2$. Let $\epsilon > 0$ and let $\alpha = \mu_2/\lambda_2 - \epsilon$ and $\beta = \mu_2/\lambda_2 + \epsilon$. We have

$$y - (\alpha - \epsilon) x = \mu_1 u + \mu_2 v - (\alpha - \epsilon)(\lambda_1 u + \lambda_2 v)$$

$$= \left(\mu_1 - \left[\frac{\mu_1}{\lambda_2} - \epsilon\right]\lambda_1\right) u + \lambda_2 \epsilon v$$

which is an element of $\mathring{C}$, hence C because its v coordinate is strictly positive, so $\alpha x \leqslant_C y$. A similar argument shows that $y \leqslant_C \beta x$, so since ϵ may be arbitrarily small we have

$$M(y/x) = m(y/x) = \frac{\mu_2}{\lambda_2}.$$

Thus, $\omega(y/x) = 0$ and if x is comparable to y then $d(x,y) = 0$. ∎

To use these results concerning the closures of cones instead of a more detailed classification scheme based on the cones themselves, we need to know that if $\bar{C}$ is a cone then it induces the same projective metric and oscillation as C. This is the content of the next lemma.

LEMMA 4·3. *Let C be a cone in a real topological vector space V such that $\bar{C}$ is also a cone. Then for all $x \in \mathring{C}$ and $y \in V$, x dominates y with respect to C if and only if x dominates y with respect to $\bar{C}$, and in this case*

$$M(y/x; C) = M(y/x; \bar{C}),$$
$$m(y/x; C) = m(y/x; \bar{C}).$$

Proof. It is clear that if x dominates y in C then x dominates y in $\bar{C}$ and

$$M(y/x; C) \geqslant M(y/x; \bar{C}),$$
$$m(y/x; C) \leqslant m(y/x; \bar{C}).$$

Because of the identities $M(y+\lambda x/x) = M(y/x)+\lambda$ and $m(y+\lambda x/x) = m(y/x)+\lambda$, which are true for any partially ordered vector space, and the fact that $y+\lambda x \in \mathring{C}$ for sufficiently large λ, it is sufficient to prove the lemma for $y \in \mathring{C}$.

Let $\alpha = m(y/x; \bar{C})$ and $\beta = M(y/x; \bar{C})$. Since $\bar{C}$ is closed, $\alpha x \leqslant y \leqslant \beta x$, and since $y \in \mathring{C}$, α and β are both strictly positive. Writing $\alpha x \leqslant y \leqslant \beta x$ in terms of the cone, we have that $y - \alpha x$ and $\beta x - y$ are elements of $\bar{C}$. Since $y \in \mathring{C}$ and $y - \alpha x \in \bar{C}$, we have by Lemma 3·8 that for $t \in (0,1)$,

$$(1-t)\,y + t(y-\alpha x) \in \mathring{C}$$

which implies that $\alpha t x \leqslant_C y$ for all $t \in (0,1)$, so $m(y/x; C) \geqslant \alpha = m(y/x; \bar{C})$. A similar argument shows that $M(y/x; C) \leqslant M(y/x; \bar{C})$. ∎

COROLLARY 4·4. *Let C be a cone in a two-dimensional real vector space V such that $\bar{C}$ is also a cone. Then for all $x \in C$ and $y \in V$, x dominates y with respect to C if and only if x dominates y with respect to $\bar{C}$, and in this case*

$$M(y/x; C) = M(y/x; \bar{C}),$$
$$m(y/x; C) = m(y/x; \bar{C}).$$

Proof. If $x \in \mathring{C}$, this is a special case of the previous result. If $x = 0$, then it is trivially true. As in the previous result, it is sufficient to prove this for y comparable to x, by adding a sufficiently large multiple of x to y. If $x \in \partial C \backslash \{0\}$, then by the representation in Theorem 4·1, we have that for some $v \in \bar{C}$,

$$\bar{C} = \{\lambda x + \mu v : \lambda, \mu \geqslant 0\},$$
$$\mathring{C} = \{\lambda x + \mu v : \lambda, \mu > 0\}.$$

Suppose x dominates y in $\bar{C}$, so $\alpha x \leqslant_{\bar{C}} y \leqslant_{\bar{C}} \beta x$. Let $y = \lambda x + \mu v$, then by the representation of $\bar{C}$ we have

$$\alpha \leqslant \lambda \leqslant \beta, \quad 0 \leqslant \mu \leqslant 0,$$

so $\mu = 0$. Thus, y is a scalar multiple of x and $M(y/x) = m(y/x)$, regardless of the cone.

Finally, if x dominates y in C, then it dominates y in $\bar{C}$ since $\bar{C} \supseteq C$, so by the previous part, $m(y/x) = M(y/x)$ in either cone. ∎

COROLLARY 4·5. *In the notation of Theorem* 3·1, *if* $\bar{C}$ *and* $\bar{D}$ *are cones and* $\dim(V) = \dim(W) = \dim(C) = \dim(D) = 2$, *then*

$$k(L;C,D) = k(L;\bar{C},\bar{D}),$$
$$N(L;C,D) = N(L;\bar{C},\bar{D}),$$
$$\Delta(L;C,D) = \Delta(L;\bar{C},\bar{D}).$$

Proof. This result follows easily from 4·4. ∎

LEMMA 4·6. *To prove Theorem* 3·5, *it is sufficient to prove it in the case that* $\dim(V) = \dim(W) = \dim(C) = \dim(D) = 2$, $\bar{C}$ *and* $\bar{D}$ *are cones and* L *is injective.*

Proof. We already know that it is sufficient to prove Theorem 3·5 when $\dim(V) = \dim(W) = \dim(C) = \dim(D) = 2$ and L is injective. If $\bar{C}$ or $\bar{D}$ is not a cone, then it is immediate from Corollary 4·2 and the fact that A does not increase either d or ω that $k(A;C,D)$, $N(A;C,D)$ and $\Delta(A;C,D)$ are all zero, so Theorem 3·5 is trivially true. ∎

LEMMA 4·7. *To prove Theorem* 3·5, *it is sufficient to prove it in the case that* $V = W = \mathbb{R}^2$, $C = D = \{x \in \mathbb{R}^2 : x_1, x_2 \geqslant 0\}$ *and* L *is represented with respect to the standard basis by the matrix*

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

where a, b, c, $d > 0$ *and* $ad-bc \neq 0$.

Proof. We know by Lemma 4·6, that to prove Theorem 3·5, it is sufficient to prove it in the case that $\dim(V) = \dim(W) = \dim(C) = \dim(D) = 2$, $\bar{C}$ and $\bar{D}$ are cones and $L: V \to W$ is injective with $L(C) \subseteq D$.

Corollary 4·5 implies that we may replace C by $\bar{C}$ and D by $\bar{D}$ without affecting k, N or Δ, and hence assume that C and D are closed.

Theorem 4·1 implies that there are linearly independent vectors u, $v \in C$ and $\tilde{u}$, $\tilde{v} \in D$ such that

$$C = \{\alpha u + \beta v : \alpha, \beta \geqslant 0\},$$
$$D = \{\alpha \tilde{u} + \beta \tilde{v} : \alpha, \beta \geqslant 0\}.$$

Define $S: \mathbb{R}^2 \to V$ and $T: W \to \mathbb{R}^2$ by

$$S(\alpha, \beta) = \alpha u + \beta v,$$
$$T(\alpha \tilde{u} + \beta \tilde{v}) = (\alpha, \beta).$$

We have that S and T are injective linear maps with $S(K) = C$ and $T(D) = K$, where K is the standard cone in $\mathbb{R}^2$. By Lemma 3·12, it suffices to prove Theorem 3·5 with L replaced by $L_1 = TLS$, $C = D = K$ and $V = W = \mathbb{R}^2$. Let L_1 be represented with respect to the standard basis by the matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

so $L_1(1,0) = (a,c)$ and $L(0,1) = (b,d)$. Because $L_1(K) \subseteq K$, a, b, c and d are non-negative; because L is injective, $ad-bc \neq 0$. We are also assuming that $\Delta(L_1;K,K) < \infty$, so by Proposition 2·9, (a,c) and (b,d) are comparable in K. A little thought shows that this is true only if either $a = b = 0$, $c = d = 0$ or all of a, b, c and d are positive. Either of the first two cases contradict the hypothesis that $ad - bc \neq 0$, so we must have $a, b, c, d > 0$. ∎

5. *Completion of the proof of the Birkhoff–Hopf theorem*

We know from Lemma 4·7 that to prove Theorem 3·5, it is sufficient to prove it in the case that $V = W = \mathbb{R}^2$, $C = D = \{x \in \mathbb{R}^2 : x_1, x_2 \geqslant 0\}$ and L is represented with respect to the standard basis by the matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

where $a, b, c, d > 0$ and $ad - bc \neq 0$.

To prove the theorem for this restricted case, we shall use the simple formulae for ω and d in $\mathbb{R}^2$ to convert the problem into a simple exercise in calculus.

Before we start, we shall use one final reduction. This is not strictly necessary, but it substantially simplifies the calculus.

Lemma 5·1. *To prove Theorem* 3·5, *it is sufficient to prove it in the case that* $V = W = \mathbb{R}^2$, $C = D = \{x \in \mathbb{R}^2 : x_1, x_2 \geqslant 0\}$ *and* L *is represented with respect to the standard basis by the matrix*

$$A = \begin{bmatrix} \alpha & 1 \\ 1 & \alpha \end{bmatrix},$$

where $\alpha > 1$.

Proof. Bearing in mind Lemma 4·7, all we need to show is that, with the given spaces and cones, then given a matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

there exists a matrix

$$A' = \begin{bmatrix} \alpha & 1 \\ 1 & \alpha \end{bmatrix},$$

with $k(A') = k(A)$, $N(A') = N(A)$ and $\Delta(A') = \Delta(A)$.

We shall construct this matrix by multiplication on the left and right by matrices which are bijections on the cone. According to Lemma 3·12, this does not affect N, k or Δ.

We begin by finding positive diagonal matrices D_1 and D_2 such that $D_1 A D_2$ is doubly stochastic, that is that its row and column sums are all 1. Let

$$D_1 = \begin{bmatrix} x_1 & 0 \\ 0 & x_2 \end{bmatrix}, \quad D_2 = \begin{bmatrix} y_1 & 0 \\ 0 & y_2 \end{bmatrix}$$

so

$$D_1 A D_2 = \begin{bmatrix} a x_1 y_1 & b x_1 y_2 \\ c x_2 y_1 & d x_2 y_2 \end{bmatrix}.$$

Now, for $0 \leqslant x_1 \leqslant 1$, let $x_2 = 1 - x_1$, $y_1 = 1/(a x_1 + c x_2)$ and $y_2 = 1/(b x_1 + d x_2)$. This makes the columns of D_1AD_2 sum to 1; to make the rows also sum to 1, it is sufficient to have

$$\frac{ax_1}{ax_1+cx_2}+\frac{bx_1}{bx_1+dx_2}=1.$$

Now let

$$\theta(x_1)=\frac{ax_1}{ax_1+c(1-x_1)}+\frac{bx_1}{bx_1+d(1-x_1)}.$$

Since $\theta(0)=0$ and $\theta(1)=2$, there exists $x_1\in(0,1)$ such that $\theta(x_1)=1$. With this value of x_1, and x_2, y_1 and y_2 as defined above, D_1 and D_2 are positive and D_1AD_2 is positive and doubly stochastic, so

$$D_1AD_2=\begin{bmatrix}\beta & 1-\beta\\ 1-\beta & \beta\end{bmatrix}.$$

Now, if $\det(A)>0$, let P be the identity matrix, and if $\det(A)<0$ let P be the permutation matrix

$$\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$$

(recall that $\det(A)\neq 0$ by hypothesis), so

$$PD_1AD_2=\begin{bmatrix}\gamma & 1-\gamma\\ 1-\gamma & \gamma\end{bmatrix},$$

where $\gamma>1/2$. Finally, let D be $1/(1-\gamma)$ times the identity matrix, so

$$A'=PD_1AD_2D=\begin{bmatrix}\alpha & 1\\ 1 & \alpha\end{bmatrix},$$

where $\alpha>1$.

Since A' was obtained from A by multiplication on the left and right by matrices which are bijections on the cone, A' has the same projective diameter, oscillation ratio and contraction ratio as A. ∎

Remark 5·2. If A is any $n\times n$ matrix, all of whose entries are positive, then there exist positive diagonal matrices D_1 and D_2 such that D_1AD_2 is doubly stochastic, and such matrices are unique up to scalar multiplication. Such results are called DAD theorems; references to the relevant literature may be found in section 4 of [**22**], and in [**6**] and [**24**]. Clearly, the key result above is a direct proof of the DAD theorem for 2×2 matrices; conversely, it is a special case of results in section 4 of [**22**] that Theorem 3·5 can be used to give an easy direct proof of the above-mentioned DAD theorem in the $n\times n$ case. Thus, there is a close connection between DAD theorems and Theorem 3·5.

We may now give the simple proof of the Birkhoff–Hopf Theorem for the case of positive matrices of the type given in Lemma 5·1.

THEOREM 5·3. *Let A be the matrix*

$$\begin{bmatrix}\alpha & 1\\ 1 & \alpha\end{bmatrix},$$

where $\alpha>1$, and consider A acting on $\mathbb{R}^2$, partially ordered by the usual cone. Then

$$N(A)=k(A)=\tanh\tfrac{1}{4}\Delta(A).$$

Proof. We have

$$N(A) = \sup_{\substack{x,y\in C \\ y \text{ dominates } x \\ \omega(x/y)\neq 0}} \frac{\omega(Ax/Ay)}{\omega(x/y)}.$$

Since $\omega(Ax/Ay)/\omega(x/y)$ is independent of positive scalar multiples of either x or y, we may take a cross-section through the cone. It is convenient to consider $x = (1, s)$ and $y = (1, t)$ for non-negative s and t. For $\omega(x/y)$ to exist and be non-zero, we must have $s, t > 0$ and $s \neq t$. Thus,

$$N(A) = \sup_{s,t>0,\, s\neq t} \frac{\omega(Ax/Ay)}{\omega(x/y)}.$$

Now, if we use the standard formula for ω in $\mathbb{R}^2$,

$$N(A) = \sup_{\substack{s,t>0 \\ s\neq t}} \frac{\left|\dfrac{\alpha+s}{\alpha+t} - \dfrac{1+\alpha s}{1+\alpha t}\right|}{\left|1-\dfrac{s}{t}\right|}$$

$$= \sup_{\substack{s,t>0 \\ s\neq t}} \frac{(\alpha^2-1)\,t}{(\alpha+t)(\alpha t+1)}$$

$$= \sup_{t>0} \phi(t)$$

where

$$\phi(t) = \frac{(\alpha^2-1)\,t}{(\alpha+t)(\alpha t+1)}.$$

The maximization of ϕ over $(0, \infty)$ is a simple problem: ϕ is non-negative, its only stationary point is at 1 and its limits at 0 and ∞ are both zero; it follows that its supremum is attained at 1 and is equal to

$$\phi(1) = \frac{\alpha^2-1}{(\alpha+1)^2} = \frac{\alpha-1}{\alpha+1}.$$

A similar approach to the contraction ratio gives an apparently more difficult maximization problem. We have

$$k(A) = \sup_{\substack{s,t>0 \\ s\neq t}} \left| \frac{\log \dfrac{(\alpha+s)(1+\alpha t)}{(\alpha+t)(1+\alpha s)}}{\log \dfrac{s}{t}} \right|$$

$$= \sup_{\substack{s,t>0 \\ s\neq t}} |\psi(s,t)|,$$

where

$$\psi(s,t) = \frac{\log \dfrac{(\alpha+s)(1+\alpha t)}{(\alpha+t)(1+\alpha s)}}{\log \dfrac{s}{t}}.$$

It is convenient to write

$$\psi(s,t) = \frac{f(s)-f(t)}{\log s - \log t},$$

where

$$f(t) = \log\frac{\alpha+t}{1+\alpha t}.$$

Using the generalized mean value theorem, we have that for $0 < s < t$ there exists τ with $s \leqslant \tau \leqslant t$ and

$$\begin{aligned}\psi(s,t) &= \frac{f(s)-f(t)}{\log s-\log t}\\ &= f'(\tau)\,\tau\\ &= \frac{(1-\alpha^2)\,\tau}{(\alpha+\tau)(1+\alpha\tau)}\\ &= -\phi(\tau).\end{aligned}$$

It is immediate from this that $\sup|\psi| \leqslant \sup\phi$, both suprema being taken over the domains of the functions. To show that $\sup|\psi| \geqslant \sup\phi$, fix $t > 0$ and choose an arbitrary positive ϵ. By the mean value theorem argument above, $-\psi(t-\epsilon, t+\epsilon) = \phi(\tau)$ for some $\tau \in [t-\epsilon, t+\epsilon]$. Since ϕ is continuous at t and ϵ may be made arbitrarily small, this shows that $|\psi|$ attains values arbitrarily close to $\phi(t)$ for any given t. It follows that $\sup|\psi| \geqslant \sup\phi$.

We thus have

$$N(A) = k(A) = \frac{\alpha-1}{\alpha+1}.$$

It remains to be shown that $\tanh\frac{1}{4}\Delta(A) = (\alpha-1)/(\alpha+1)$.

Let $e_1 = (1,0)$ and $e_2 = (0,1)$ be the standard basis vectors for $\mathbb{R}^2$. $A(C)$ is given by the set of all non-negative linear combinations of Ae_1 and Ae_2; it follows from Proposition 2·9 that

$$\Delta(A) = \operatorname{diam}(A(C)) = d(Ae_1, Ae_2).$$

Since $Ae_1 = (\alpha, 1)$ and $Ae_2 = (1,\alpha)$, we have $d(Ae_1, Ae_2) = \log\alpha^2 = 2\log\alpha$.

Now, since $\Delta(A) = 2\log\alpha$, $\alpha = \exp(\Delta(A)/2)$ and

$$N(A) = k(A) = \frac{\alpha-1}{\alpha+1} = \frac{\exp(\Delta(A)/2)-1}{\exp(\Delta(A)/2)+1} = \tanh\tfrac{1}{4}\Delta(A). \quad \blacksquare$$

Finally, we show that if the projective diameter is infinite then the oscillation and contraction ratios are both equal to 1.

THEOREM 5·4. *Let V and W be real vector spaces partially ordered by cones C and D, respectively. Suppose that L is a linear operator from V to W with $L(C) \subseteq D$ and $\Delta(L; C, D) = \infty$. Then*

$$N(L; C, D) = k(L; C, D) = 1.$$

Proof. Given $M > 0$, there exist x and y such that $d(Lx, Ly) > M$. Consider the two-dimensional subspaces V' generated by x and y and W' generated by Lx and Ly (Lx

and Ly are linearly independent since $d(Lx, Ly) > 0$; this implies that x and y are also linearly independent).

We place a partial ordering on W' by means of the cone D and on V' by means of the cone C' generated by x and y, so

$$C' = \{\lambda x + \mu y : \lambda, \mu \geqslant 0\}.$$

Since C' is a subcone of C, $d(u, v; C') \geqslant d(u, v; C)$.

Now, regarding L as a map from V' ordered by C' to W' ordered by D, L has projective diameter $d(Lx, Ly)$ (by Proposition 2·9), so we may conclude from the result for the finite projective diameter case that given $\epsilon > 0$ there exist $u, v \in C'$ with

$$\frac{d(Lu, Lv; D)}{d(u, v; C')} \geqslant \tanh \tfrac{1}{4} d(Lx, Ly; D) - \epsilon.$$

Now, $d(Lx, Ly; D) \geqslant M$ and $d(u, v; C') \geqslant d(u, v; C)$, so

$$\frac{d(Lu, Lv; D)}{d(u, v; C)} \geqslant \tanh \tfrac{1}{4} M - \epsilon.$$

Thus, by choosing large M and small ϵ, we may find u and v such that $d(Lu, Lv; D)/d(u, v; C)$ is arbitrarily close to 1. Since we know *a priori* that $k(L) \leqslant 1$, we conclude that $k(L) = 1$.

A similar proof shows that $N(L) = 1$. ∎

6. *Positive matrices and integral operators*

In the special case that $V = \mathbb{R}^n$, $W = \mathbb{R}^m$, $C = \{x \in \mathbb{R}^n : x_i \geqslant 0 \text{ for all } i\}$ and $D = \{x \in \mathbb{R}^m : x_i \geqslant 0 \text{ for all } i\}$, Theorem 3·5 is usually stated in a sharper form with an explicit formula for $\Delta(L; C, D)$ in terms of the matrix representation of L. To establish this formula, it will be useful first to prove more general results.

If we take C as in Remark 2·10 and $L: C \to C$ to be the identity map, then $\Delta(L; C, C) = 0$ and $L(C)$ contains non-zero elements which are not comparable in C. If, however, we impose mild additional conditions then Proposition 2·9 implies that (in the notation of Theorem 3·5) all non-zero elements of $L(C)$ are comparable. A slightly less general version of this result has been obtained by Eveson [**11**].

Lemma 6·1. *Let the notation and hypotheses be as in Theorem 3·5. Assume that if u and v are any non-zero elements of $L(C)$ such that u dominates v in D, then $M(v/u) > 0$ (this will be the case if $\overline{D(u, v)}$ is a cone whenever u and v are non-zero elements of $L(C)$ such that u dominates v in D). Then it follows that all non-zero elements of $L(C)$ are comparable in D.*

Proof. Let $S = L(C) \subseteq D$, so S is convex and $\operatorname{diam}(S; D) < \infty$. It now follows from Proposition 2·9 that any two non-zero elements of S are comparable in D. ∎

We now return to the case where $V = \mathbb{R}^n$, $W = \mathbb{R}^m$ and C and D are the cones of non-negative vectors in V and W, respectively. If we consider the elements of $\mathbb{R}^n$ as column vectors and if $A = (a_{ij})$ is an $m \times n$ matrix, then A defines a linear map on $\mathbb{R}^n$ and if $a_{ij} \geqslant 0$ for $1 \leqslant i,j \leqslant n$, then $A(C) \subseteq D$. We let e_j $(1 \leqslant j \leqslant n)$ denote the jth standard basis vector for $\mathbb{R}^n$, so Ae_j is the jth column of the matrix A. The reader may easily verify that if x and y are elements of $\mathring{C}$ then

$$d(x,y;C) = \log \max_{i,j} \frac{x_i y_j}{x_j y_i}, \tag{7}$$

with a corresponding formula for the cone D.

THEOREM 6·2. *Let $V = \mathbb{R}^n$, $C = \{x \in \mathbb{R}^n : x_i \geqslant 0$ for $1 \leqslant i \leqslant n\}$, $W = \mathbb{R}^m$ and $D = \{x \in \mathbb{R}^m : x_i \geqslant 0$ for $1 \leqslant i \leqslant m\}$. Let A be an $m \times n$ matrix (a_{ij}) such that $a_{ij} \geqslant 0$ for $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$, so $A(C) \subseteq D$. If $\{e_i : 1 \leqslant i \leqslant n\}$ is the standard orthonormal basis for $\mathbb{R}^n$, assume that there exists a set $J \subseteq \{1, 2, \ldots, n\}$ such that Ae_i and Ae_j are non-zero and comparable in D for each $i, j \in J$ and $Ae_i = 0$ for all $i \notin J$. Then it follows that $\Delta(A) = \Delta(A; C, D)$ is finite and*

$$\Delta(A) = \max_{i,j \in J} d(Ae_i, Ae_j). \tag{8}$$

If, moreover, $a_{ij} > 0$ for all $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$ then

$$\Delta(A) = \max_{1 \leqslant i,j \leqslant n} d(Ae_i, Ae_j) = \log \max_{\substack{1 \leqslant i,j \leqslant n \\ 1 \leqslant p,q \leqslant m}} \frac{a_{pi} a_{qj}}{a_{pj} a_{qi}}. \tag{9}$$

Proof. Let $T = \{Ae_i : i \in J\}$ and observe that

$$A(C) = \{tx : t \geqslant 0 \text{ and } x \in \operatorname{co}(T)\}.$$

By using parts (*a*) and (*b*) of Proposition 2·9, we see that

$$\Delta(A) = \operatorname{diam}(A(C); D) = \operatorname{diam}(\operatorname{co}(T); D) = \operatorname{diam}(T; D)$$

which yields (8). Equation (9) follows from this and the explicit formula for $d(Ae_i, Ae_j)$ in (7). ∎

Equation (9) is closely related to results of E. Hopf[**16**], who observed that analogous formulae hold for integral operators.

THEOREM 6·3. (See Hopf[**16**]). *Suppose S is a compact Hausdorff space, μ is a regular Borel measure on S of full support (so $\mu(G) > 0$ for any open set $G \subseteq S$) and $k : S \times S \to \mathbb{R}$ is a positive continuous function. Let $V = C(S)$, the Banach space of continuous, real-valued functions on S and let C be the cone of non-negative functions in V. Define $A : V \to V$ by*

$$(Ax)(s) = \int_S k(s,t)\, x(t)\, \mu(dt) \tag{10}$$

Then it follows that

$$\Delta(A) = \Delta(A; C, C) = \log \max_{s,t,u,v} \frac{k(s,t)\, k(u,v)}{k(s,v)\, k(u,t)}. \tag{11}$$

Proof. If x and y are non-zero elements of C we have

$$M(Ay/Ax) = \max_s \frac{(Ay)(s)}{(Ax)(s)} = \max_s \frac{\int_S k(s,\tau)\,y(\tau)\,\mu(d\tau)}{\int_S k(s,\nu)\,x(\nu)\,\mu(d\nu)},$$

$$(m(Ay/Ax))^{-1} = M(Ax/Ay) = \max_u \frac{\int_S k(u,\nu)\,x(\nu)\,\mu(d\nu)}{\int_S k(u,\tau)\,y(\tau)\,\mu(d\tau)}.$$

Thus we obtain that

$$M(Ay/Ax)M(Ax/Ay) = \max_{s,u} \frac{\iint k(s,\tau)\,k(u,\nu)\,x(\nu)\,y(\tau)\,\mu(d\nu)\,\mu(d\tau)}{\iint k(s,\nu)\,k(u,\tau)\,x(\nu)\,y(\tau)\,\mu(d\nu)\,\mu(d\tau)}.$$

From this formula, we easily conclude that

$$\begin{aligned}\chi A\,; C, C) &= \chi(A)\\ &= \inf\{\lambda \geqslant 0: \iint [k(s,\tau)\,k(u,\nu) - \lambda^2 k(s,\nu)\,k(u,\tau)]\,x(\nu)\,y(\tau)\,\mu(d\nu)\,\mu(d\tau) \leqslant 0\\ &\qquad \text{for all } x, y \in C\backslash\{0\} \text{ and } s, u \in S\}.\end{aligned}$$

If we define Λ by

$$\Lambda^2 = \max_{s,\tau,u,\nu} \frac{k(s,\tau)\,k(u,\nu)}{k(s,\nu)\,k(u,\tau)},$$

we obviously have $\chi(A) \leqslant \Lambda$. If $0 < \lambda < \Lambda$, we have

$$k(s,\tau)\,k(u,\nu) - \lambda^2 k(s,\nu)\,k(u,\tau) > 0$$

for $(s,u) \in U_1 \times W_1$ and $(\nu,\tau) \in U_2 \times W_2$, where U_j and W_j are open sets. There exist continuous functions $x, y \in C\backslash\{0\}$ which vanish outside U_2 and W_2 respectively. If $(s,u) \in U_1 \times W_1$, we find that

$$M(Ay/Ax)M(Ax/Ay) \geqslant \frac{(Ay)(s)(Ax)(u)}{(Ax)(s)(Ay)(u)} > \lambda^2$$

so $\chi(A) > \lambda$. It follows that $\chi(A) = \Lambda$ and that (11) holds. ∎

Remark 6·4. If μ is not of full support, one obtains

$$\Delta(A) \leqslant \log \max_{s,t,u,v} \frac{k(s,t)\,k(u,v)}{k(s,v)\,k(u,t)},$$

but equality need not hold, as one may see by taking $\mu = \delta$, the Dirac delta measure.

If $k \in L^1(S \times S)$ is a non-negative function such that

$$\operatorname{ess.sup}_s \int k(s,t)\,\mu(dt) < \infty,$$

then k defines a bounded integral operator $A: E = L^\infty(S) \to E$ by (10). If C is the cone of non-negative functions in E, the argument above shows that

$$\chi(A) \leqslant \inf\{\lambda \geqslant 0: k(s,\tau)\,k(u,\nu) - \lambda^2 k(s,\nu)\,k(u,\tau) \leqslant 0 \text{ for almost all } s, u, \nu, \tau\}.$$

However, our previous arguments do not prove that equality holds in (11).

We note that in the generality of Remark 6·4 it is easy to find functions $k \in L^{\infty}(S \times S)$ such that $A : L^{\infty} \to L^{\infty}$ is not a compact map: see remark 5·6 in section 5 of [6].

REFERENCES

[1] F. L. BAUER. An elementary proof of the Hopf inequality for positive operators. *Numerische Math.* 7 (1965), 331–337.

[2] G. BIRKHOFF. Extensions of Jentzsch's Theorem. *Trans. Amer. Math. Soc.* **85** (1957), 219–226.

[3] G. BIRKHOFF. Uniformly semi-primitive multiplicative processes. *Trans. Amer. Math. Soc.* **104** (1962), 37–51.

[4] G. BIRKHOFF and L. KOTIN. Asymptotic behaviour of solutions of first order linear differential-delay equations. *J. Math. Anal. App.* **13** (1966), 8–18.

[5] G. BIRKHOFF. Integro-differential delay equations of positive type. *J. Differential Equations* **2** (1966), 320–327.

[6] J. M. BORWEIN, A. S. LEWIS and R. D. NUSSBAUM. Entropy minimization, DAD problems and doubly stochastic kernels. *J. Functional Analysis*, to appear.

[7] P. J. BUSHELL. Hilbert's projective metric and positive contraction mappings in a Banach space. *Arch. Rational Mech. Anal.* **52** (1973), 330–338.

[8] P. J. BUSHELL. On the projective contraction ratio for positive linear mappings. *J. London Math. Soc.* **6** (1973), 256–258.

[9] P. J. BUSHELL. The Cayley–Hilbert metric and positive operators. *Lin. Alg. Appl.* **84** (1986), 271–280.

[10] J. E. COHEN. Ergodic theorems in demography. *Bull. Amer. Math. Soc.* **1** (1979), 275–295.

[11] S. P. EVESON. Hilbert's projective metric and the spectral properties of positive linear operators. *Proc. London Math. Soc.*, to appear.

[12] S. P. EVESON. *Theory and applications of Hilbert's projective metric to linear and nonlinear problems in positive operator theory.* Ph.D. thesis, University of Sussex, 1991.

[13] S. P. EVESON and R. D. NUSSBAUM. Applications of the Birkhoff–Hopf Theorem to the spectral theory of positive linear operators. *Math. Proc. Camb. Phil. Soc.*, to appear.

[14] T. FUJIMOTO and U. KRAUSE. Asymptotic properties of inhomogeneous iterations of nonlinear operators. *SIAM J. Math. Anal.* **19** (1988), 841–853.

[15] M. GOLUBITSKY, E. KEELER and M. ROTHSCHILD. Convergence of the age structure: Application of the projective metric. *Theoretic. Population Biol.* **7** (1975), 84–93.

[16] E. HOPF. An inequality for positive integral operators. *J. Math. Mech.* **12** (1963), 683–692.

[17] E. HOPF. Remarks on my paper 'An inequality for positive integral operators'. *J. Math. Mech.* **12** (1963), 889–892.

[18] M. A. KRASNOSEL'SKII, JE. A. LIFSHITS and A. V. SOBOLEV. *Positive linear systems: the method of positive operators.* Sigma Series in Applied Mathematics, vol. 5 (Heldermann Verlag, 1989).

[19] M. A. KRASNOSEL'SKII and A. V. SOBOLEV. Spectral clearance of a focusing operator. *Funct. Anal. Appl.* **17** (1983), 58–59.

[20] M. G. KREIN and M. A. RUTMAN. Linear operators leaving invariant a cone in a Banach space. *Transl. Amer. Math. Soc.* **10** (1962), 199–325 (English translation by M. M. Day).

[21] R. D. NUSSBAUM. *Hilbert's projective metric and iterated nonlinear maps.* Memoirs of the American Math. Soc. **75** (1988), 391.

[22] R. D. NUSSBAUM. *Hilbert's projective metric and iterated nonlinear maps II.* Memoirs of the American Math. Soc. **79** (1989), 401.

[23] R. D. NUSSBAUM. Some nonlinear weak ergodic theorems. *SIAM J. Math. Anal.* **21** (1990), 436–460.

[24] R. D. NUSSBAUM. Entropy minimization. Hilbert's projective metric and scaling integral kernels. *J. Functional Analysis*, to appear.

[25] A. M. OSTROWSKI. On positive matrices. *Math. Ann.* **150** (1963), 276–284.

[26] A. M. OSTROWSKI. Positive matrices and functional analysis in *Recent Advances in Matrix Theory* (University of Wisconsin Press, 1964) pp. 81–101.

[27] W. RUDIN. *Functional analysis* (McGraw-Hill, 1973).

[28] H. H. SCHAEFER. *Topological vector spaces* (Macmillan, 1966).

[29] K. WYSOCKI. Behaviour of directions of solutions of differential equations. *Differential and Integral Equations* **5** (1992), 281–305.

[30] K. WYSOCKI. Ergodic theorems for solutions of differential equations. *SIAM J. Math. Analysis* (1993), to appear.

[31] P. P. ZABREIKO, M. A. KRASNOSEL'SKII and YU. V. POKORNYI. On a class of positive linear operators. *Functional Analysis and its Applications* **5** (1972), 272–279.

[32] A. D. ZIEBUR. New directions in linear differential equations. *SIAM Review* **21** (1979), 57–70.