EXTENSIONS OF JENTZSCH'S THEOREM

BY

GARRETT BIRKHOFF

1. Introduction. In this paper, the projective metric of Hilbert(1) is applied to prove various extensions of Jentzsch's theorem on integral equations(2) with positive kernels. In particular, it is shown that Jentzsch's theorem reduces to the Picard fixpoint theorem(3), relative to this projective metric.

A natural setting for generalizing Jentzsch's theorem seems to be provided by the theory of vector lattices(4). A bounded linear transformation $P$ of a vector lattice $L$ into itself will be called uniformly positive if, for some fixed $e > 0$ in $L$ and finite real number $K$, independent of $f$, we have

\[ \lambda e \leq fP \leq K\lambda e \quad \text{for any } f > 0 \text{ and some } \lambda = \lambda(f) > 0. \]

Theorem 3 below shows that Jentzsch's theorem applies, in generalized form, to any such operator $P$.

The method is also applied to various other cases: in §5, to a class of integro-functional equations to which the usual proof would not be applicable; in §8, to a class of semigroups including various multiplicative processes(6).

2. Projective metrics on line. For convenient reference, we derive some basic formulas regarding the effect of projective transformations on projective metrics. In homogeneous coordinates, the first positive quadrant joins $(0, 1)$ with $(1, 0)$ by "points" $(f_1, f_2)$. This is mapped onto the hyperbolic line $-\infty < u < +\infty$ by the correspondence $Ln(f_2/f_1) = u$. We define

\[ \theta(f, g) = \left| Ln v - Ln u \right| = \left| Ln \left( \frac{f_2g_1}{f_1g_2} \right) \right|. \]

Since $f_2g_1/f_1g_2$ is the cross-ratio $R(f_2/f_1, g_2/g_1; 0, \infty)$, $\theta(f, g)$ is invariant under all projective transformations mapping the interval $0 < f_2/f_1 < +\infty$ onto itself.

We next consider a general projective transformation
which maps $0 < x < \infty$ onto a proper subinterval of $0 < y < \infty$. Without loss of generality, since $y \to 1/y$ is an isometry for $\theta(y, y')$, we can assume that order is preserved. That is, we can assume $x > x'$ implies $xP \geq x'P$. Since $OP = b/d$, clearly $b$ and $d$ have the same sign; since $xP = 0$ has no positive solution, $b$ and $a$ have the same sign; since $xP = \infty$ has no positive solution, $d$ and $c$ have the same sign. Hence we can assume $a, b, c, d$ positive in (3). Furthermore, since

$$y' = \frac{dy}{dx} = \frac{(ad - bc)}{(cx + d)^2},$$

we can assume $ad > bc$.

We now consider the ratio of hyperbolic distance differentials $d\theta(y)/d\theta(x) = xdy/ydx$. By (3) and (4), this is $d\theta(y)/d\theta(x) = (ad - bc)x/(ax+b)(cx+d)$. By the differential calculus, the maximum of this (i.e., the minimum of its reciprocal) occurs when $ac = bdx^{-2}$, or $x = (bd/ac)^{1/2}$. Substituting above and simplifying, we see that $P$ contracts all hyperbolic distances, by a factor whose supremum is

$$N(P) = \frac{\nu - 1}{\nu + 2(\nu)^{1/2} + 1},$$

where $\nu = (ad/bc) > 1$.

We call $N(P)$ the projective norm of $P$. Because of its definition, as the supremum of distance ratios

$$N(P) = \sup \left[ \frac{\theta(fP, gP)}{\theta(f, g)} \right],$$

we have immediately

$$N(PP') \leq N(P)N(P').$$

Also, if $\lambda = \ln (ad/bc)$ is the length of the segment onto which $P$ maps the first quadrant, then by (5), we have

$$N(P) = \frac{\nu^{1/4} - \nu^{1/4}}{\nu^{1/4} + \nu^{1/4}} = \frac{2 \sinh (\lambda/4)}{2 \cosh (\lambda/4)} = \tanh \frac{\lambda}{4}.$$

The intervention of hyperbolic functions is most appropriate!

3. Convex cones. Now let $C$ be any bounded closed convex cone of a real vector space $L$, of finite or infinite dimensions. It is convenient to make a central projection of $C$ onto its (convex) intersection $C \cap H$ with a hyperplane $H$, cutting each ray of $C$ in exactly one point; we can then discuss $C$ and $C \cap H$ interchangeably, as subspaces of projective space.

Since $C$ is a bounded closed convex set, every line intersects $C$ in a closed segment. Hence, if $f \neq g$ in $H$, the intersection of the line $l(f, g)$ with $C$ can be mapped onto the line $0 \leq x \leq \infty$ of $\S 2$ so that $fA < gA$ by a unique affine transformation $A$. We define
If $f$ or $g$ is a boundary point, $\theta(f, g; C) = \infty$. We call $\theta(f, g; C)$ the projective metric associated with $C$.

The following result is well-known.

**Lemma 1.** For any $a \in C$, the set $A$ of rays $f \in C$ satisfying $\theta(a, f; C) < +\infty$ is a metric space relative to the distance $\theta(f, g; C)$.

It is well known that, if $C \cap H$ is an ellipsoid, then $\theta(f, g; C)$ makes $C$ into a hyperbolic geometry. It seems not to have been observed, however, that the following example leads to the Perron-Frobenius theory of positive matrices.

**Example 1.** Let $C$ be the cone $R$ of “positive” $f \neq 0$ satisfying $f_1 \geq 0, \ldots, f_n \geq 0$ in real $n$-space, and let $H$ be the hyperplane $\sum f_i = 1$. In this case, the disconnected components of $C \cap H$ are the interiors of its cells, where $C \cap H$ is regarded as a simplex.

The theorem of Jentzsch can be deduced very simply, as we shall see below, from the following special case.

**Example 2.** Let $L$ be the space of continuous functions, in the usual Banach lattice norm $\|f\| = \sup |f(x)|$, and let $C$ be the cone $L^+$ of non-negative functions. Then the functions which are identically positive form a connected component under $\theta(f, g; L^+)$.

Similarly, generalizations of Jentzsch’s theorem can be deduced by considering other special cases, such as the following.

**Example 3.** Let $B$ be the Banach space of bounded measurable functions on the unit interval, with the norm $\|f\| = \sup |f(x)|$. For any positive constant $M$, let $C$ be the cone of functions satisfying

$$0 < \sup f(x) \leq M \inf f(x).$$

We omit proving that the cones in question are closed in the relevant Banach spaces, if the origin is included.

4. **Fixpoint theorem.** Let $P$ be any bounded linear transformation of a Banach space $B$, which maps a closed convex cone $C$ of $B$ into itself. The $C$-norm $N(P; C)$ of $P$ is defined as

$$N(P; C) = \sup \frac{\theta(fP, gP; C)}{\theta(f, g; C)},$$

for pairs $f, g \in C$ with finite $\theta(f, g; C)$.

**Lemma 1.** If the transform $CP$ of $C$ under $P$ has finite diameter $\Delta$ under $\theta(f, g; C)$, then

$$N(P, C) = \tanh (\Delta/4) < 1.$$
Proof. If \( \theta(f, g; C) < +\infty \), then \( f \) and \( g \) lie on a segment \( s(a, b) \) of \( C \). The image segment \( s(aP, bP) \leq CP \); hence \( \theta(aP, bP; C) \leq \Delta \). By (7), we infer

\[
\frac{\theta(fP, gP; C)}{\theta(f, g; C)} \leq \tanh (\Delta/4).
\]

Hence, by (9), \( N(P, C) \leq \tanh (\Delta/4) \). To show that equality holds, we take a sequence of inverse images \( f^n, g^n \) of suitable nearby pairs of points on segments \( s(c^n, d^n) \) of lengths \( \Delta - 2^{-n} \) or more, and use (7) again.

If \( CP \) has infinite diameter, then similar considerations show that \( N(P, C) = 1 \). (The fact that \( N(P, C) \leq 1 \) is immediate from (7).)

Theorem 1 (Projective contraction theorem). Let \( N(P^r; C) < 1 \) for some \( r \), and let \( C \) be complete relative to \( \theta(f, g; C) \). Then, for any \( f \in C \), the sequence of \( fP^n \) converges geometrically to a unique fixpoint (characteristic ray) \( c \in C \).

Proof. If \( N(P^r; C) < 1 \), then \( CPr \) has a finite hyperbolic diameter, by what we have just seen. Hence \( \theta(fPr, fPr+1; C) < +\infty \). More generally, if \( q > 0 \) is the integral part of \( (n/r) \), then (writing \( \theta(f, g; C) \) as \( \theta(f, g) \))

\[
\theta(fP^n, fP^{n+1}) \leq N(P^r; C)^q \theta(fP^r, fP^{r+1}).
\]

Hence, as in the proof of Picard's Fixpoint Theorem(3), \( \{fP^n\} \) is a Cauchy sequence. By the assumption of completeness, the Cauchy sequence converges to a limit \( c \in C \). Since \( P \) is bounded, \( cP = c \), and

\[
\|fP^n - c\| < K\rho^n,
\]

where \( \rho = N(P^r; C)^{1/r} \), and \( K < +\infty \). The uniqueness of \( c \) is immediate, since \( cP = c \) and \( c^*P = c^* \) imply

\[
\theta(c, c^*) = \theta(cPr, c^*Pr) \leq N(P^r, C)\theta(c, c^*),
\]

all relative to \( C \). Since \( N(P^r, C) < 1 \), this implies \( \theta(c, c^*) = 0 \).

Corollary 1. If some \( CP^* \) has finite projective diameter relative to \( C \), then the conclusion of Theorem 1 holds.

5. Applications. To apply Theorem 1, one must verify that the cone \( C \) involved is complete in the projective metric \( \theta(f, g; C) \). This is obvious in the case of Example 1, from the known(6) facts about finite-dimensional projective metrics. The cases of Examples 2–3 are also easily covered. We now give some applications of Theorem 1 based on these special examples.

If \( P \) is the linear transformation corresponding to a matrix \( \|p_{ij}\| \), with positive entries, then \( RP \) is a compact subset interior to \( R \). The rays \( L \) touching \( R \), in Example 1 of §3, are also a compact set, and the lengths \( \theta(LP) \) of their transforms vary continuously with \( L \); hence \( RP \) has finite projective diameter. We conclude, by Theorem 1, that \( P \) admits a positive eigenvector \( c \in R \), with \( cP = \gamma c \). Obviously, it is sufficient that \( P \) be non-negative, and some power \( P^r \) positive.
Again, as in Jentzsch's theorem, let $P$ be the operator defined by $[fP](x) = \int_0^1 p(x, y)f(y)dy$, $p(x, y) > 0$, with

(10) \[ 0 < I = \inf p(x, y) \leq \sup p(x, y) = KI = S. \]

Choose $L$ as in Example 2 of §3. Then, if $e(x) \equiv 1$, and $f(x) \geq 0$ with $\int f(x)dx = \phi > 0$, clearly $(I\phi)e \leq fP \leq (S\phi)e$. Hence $\theta(e, fP; L^+ ) \leq \ln K$, and Theorem 1 applies to show that $fP^n \to c$, where $cP = \gamma c$, $c \in R$.

Note that the preceding proof does not assume Fredholm's theory of integral equations. It will be generalized in Theorem 3 below.

Projective metrics can be applied flexibly\(^8\) to a variety of positive transformations. The following application is fairly typical.

**Theorem 2.** Let $p(x, y)$ satisfy (10); let $T_1, \ldots, T_n$ be any one-one Borel transformations of the unit interval onto itself, and let $a_1, \ldots, a_n$ be positive constants with sum $A$. Then the integro-functional equation

(11) \[ [fP](x) = \int_0^1 p(x, y)f(y)dy + \sum a_if(xT_i) \]

admits a unique positive characteristic function, such that

(12) \[ \int_0^1 p(x, y)c(y)dy + \sum a_ic(xT_i) = \gamma c(x). \]

**Proof.** Choose any $M > K$, and let $B$ and $C$ be defined for this $M$ as in Example 3 of §4. Then, if $f(x) \in C$ and $g(x) = [fP](x)$, we have

\[
\sup g(x) \leq KI \int f(y)dy + MA \inf f(x),
\]

\[
\inf g(x) \geq I \int f(y)dy + A \inf f(x).
\]

In view of the inequality $\int f(y)dy \geq \inf f(x)$, we infer

\[
\frac{\sup g(x)}{\inf g(x)} \leq \frac{KI + MA}{I + A} < M, \quad \text{if } K < M.
\]

Hence, the projective contraction theorem applies, and so the conclusion of Theorem 2 follows.

6. **Banach lattices.** In a different direction, one can generalize Jentzsch's

\(^8\) In this respect, the technique of projective metrics is analogous to the Leray-Schauder technique applied by E. Rothe, Amer. J. Math. vol. 66 (1944) pp. 245-254, and by M. G. Krein and M. A. Rutman, Uspehi Matematicheskikh Nauk. vol. 3 (1948) pp. 3-95, to prove other generalizations of Jentzsch’s theorem. It differs from the Leray-Schauder technique in being constructive and in not assuming complete continuity of $P$.  

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
theorem to Banach lattices. In making this generalization, the following lemma will prove convenient.

**Lemma 2.** In any vector lattice $L$, let $L^+$ denote the cone of positive elements. If $f$ and $g$ are in the same connected component of $L^+$, then they are strongly comparable in the sense that

$$\lambda f \leq g \leq R\lambda f \quad \text{and} \quad \mu g \leq f \leq R\mu f, \quad R < +\infty.$$ 

Actually, the smallest such $R = \exp\left[\theta(f, g; L^+)\right]$.

**Proof.** The plane through $f$ and $g$ intersects $L^+$ in a domain affine equivalent to a quadrant of the $(x, y)$-plane, with $g_2/g_1 \geq f_2/f_1$. In this quadrant, we easily calculate $(f_1, f_2) \leq (g_1, g_2/f_1) \leq R(f_1, f_2)$, etc. To complete the proof, consider the $(x, y)$-plane as a projective line.

**Corollary 1.** On the unit sphere of any Banach lattice $L$, we have

$$\|f - g\| \leq e^\theta - 1,$$

where $\theta = \theta(f, g; L^+)$. 

**Proof.** Suppose $\|f\| = \|g\| = 1$ in (13). Then by the monotonicity of $\|f\|$ as a function of $|f|$, $\lambda \leq 1 \leq R\lambda$. Consequently

$$\|f - g\| = \|f \cup g - f \cap g\| \leq \|R\lambda f - \lambda f\| = (R - 1)\lambda\|f\|.$$ 

Since $R = e^\theta$ and $\lambda \leq 1$, the proof is complete.

**Corollary 2.** In the metric $\theta(f, g; L^+)$, any $\theta$-connected component of the unit sphere of any Banach lattice is a complete metric space.

**Theorem 3.** Any uniformly positive bounded linear transformation $P$ of a Banach lattice $L$ into itself admits a unique positive unit vector $c$ such that

$$cP = \gamma c,$$

$\gamma > 0$. 

For any $f > 0$, $\|(fP^n/\|fP^n\|) - c\| < Mp^n$, for some finite $M$ and positive $p < 1$.

**Proof.** Choose $C$ as the set $L^+$ of positive elements. By Theorem 1 and Corollary 2 of Lemma 2, it suffices to show that $CP$ has finite projective diameter. Since $\theta(fP, gP; C) = \theta(fP/\lambda(f), gP/\lambda(g); C)$ we can assume that $\lambda(f) = \lambda(g) = 1$ in (1). Hence, if $K$ is defined by (1), the segment $(Kf - g, (K - 1)f, (K - 1)g, Kg - f)$ is in $C$. But, by the projective invariance of cross-ratios,

$$R(Kf - g, (K - 1)f, (K - 1)g, Kg - f) = R(-1, 0, K - 1, K),$$

since the two quadruples are perspective. Hence $\theta(fP, gP; C) \leq \ln K$, and the projective diameter of $CP$ is at most $\ln K$. This completes the proof.

7. **Complementary invariant subspace.** Let $P$ be again any uniformly positive linear operator on a Banach lattice $L$, and let $c$ be the associated positive characteristic vector, with positive characteristic value $\gamma$. For any positive $f > 0$, we can define $\lambda_n$ and $\mu_n$ as the largest and smallest real numbers, respectively, such that
Clearly, \(0 < \lambda_n \leq \mu_n\), for all \(n \geq 1\) (see also (16d) below). Applying \(P\) to (16a), we get \(\lambda_n \gamma^{n+1}c \leq fP^{n+1} \leq \mu_n \gamma^{n+1}c\), whence

\[
\lambda_n \leq \lambda_{n+1} \leq \mu_{n+1} = \mu_n, \text{ or } \lambda_n \uparrow \text{ and } \mu_n \downarrow.
\]

Now consider \(r_n = fP^n - \lambda_n \gamma^n c\) and \(s_n = \mu_n \gamma^n c - fP^n\). Clearly, \(0 \leq r_n, 0 \leq s_n\), and \(r_n + s_n = (\mu_n - \lambda_n) \gamma^n c\). By (1) and Lemma 2 of §6, \(\alpha_n > 0\) and \(\beta_n > 0\) exist (the case \(r_n = s_n = 0\) is trivial), such that

\[
\alpha_n c \leq r_n P \leq e^\Delta \alpha_n c \quad \text{and} \quad \beta_n c \leq s_n P \leq e^\Delta \beta_n c,
\]

whence

\[
(\alpha_n + \beta_n)c \leq (r_n + s_n) P = \gamma^{n+1}(\mu_n - \lambda_n)c \leq e^\Delta (\alpha_n + \beta_n)c.
\]

On the other hand, (16c) implies

\[
(\lambda_n \gamma^{n+1} + \alpha_n)c \leq (\lambda_n \gamma^n c + r_n P) = fP^{n+1} = (\mu_n \gamma^n c - s_n P) \leq (\mu_n \gamma^{n+1} + \beta_n)c,
\]

whence \(\lambda_{n+1} \geq \lambda_n + \gamma^{-n-1} \alpha_n, \mu_{n+1} \leq \mu_n - \gamma^{-n-1} \beta_n\). Subtracting these inequalities, and using (16d), we get

\[
\mu_{n+1} - \lambda_{n+1} \leq (\mu_n - \lambda_n) - \gamma^{-n-1}(\alpha_n + \beta_n) \leq (1 - e^{-\Delta})(\mu_n - \lambda_n).
\]

Induction on \(n\) gives \((\mu_{n+1} - \lambda_{n+1}) \leq (1 - e^{-\Delta})^n(\mu_1 - \lambda_1)\), so that \(\lambda_n \uparrow M\) and \(\mu_n \downarrow M\) for some finite \(M = M(f)\). Now, referring back to (16a), we conclude

**Lemma 3.** For each \(f > 0\), there exist positive constants \(M = M(f), K = K(f)\), and \(\rho = (1 - e^{-\Delta})\gamma < \gamma\) independent of \(f\), such that

\[
|fP^n - M\gamma^n c| \leq K\rho^n c, \quad 0 \leq \rho < \gamma.
\]

But now, \(f = f^+ + f^-\) for any \(f \in L\), where \(f^+ = f \cap 0 \geq 0\) and \(f^- = f \cap 0 \leq 0\). Writing \(fP^n = f^+P^n + f^-P^n\), we deduce the following

**Corollary.** The inequality (17) holds for each \(f \in L\), and \(M(f)\) is a positive linear functional.

**Theorem 4.** Any uniformly positive linear operator \(P\), acting on a Banach lattice \(L\), decomposes \(L\) into an invariant axis with positive basis-element (eigenvector) \(c\) and associated positive eigenvalue \(\gamma\), and a complementary invariant subspace \(S\) on which the spectral norm\(^{(9)}\) of \(P\) is at most \((1 - e^{-\Delta})\gamma < \gamma\).

**Proof.** Let \(S\) be the subspace on which \(M(f) = 0\). Then, by (17) with \(M = 0, SP \leq S\). The last conclusion also follows from Lemma 3.

8. **Multiplicative processes.** We now consider one-parameter semigroups \(\{P_t\}\) of non-negative linear operators, like those involved in multiplicative processes\(^{(6)}\). For simplicity, we shall consider only one-parameter semigroups

\[\text{L. Loomis, Abstract harmonic analysis, p. 75. The conclusion holds in any vector lattice which is complete in } \theta(f, g; L^+).\]
on Banach lattices, though the method can easily be adapted to other cases.

Accordingly, let $P_t$ ($t > 0$) map a Banach lattice $L$ linearly into itself, so that

$$f > 0 \implies fP_t > 0. \tag{18}$$

We assume the (Chapman-Kolmogorov) semigroup condition $(fP_t)P_r = fP_{t+r}$. The special case

$$f(x; t + \tau) = \int p(x; y; \tau)f(y; t)dR(y), \tag{19}$$

with $p(x, y; \tau) > 0$ for all $\tau > 0$, is typical for many applications (e.g., to multi-group diffusion).

**Theorem 5.** If, for some $t = T$, $P_T$ is uniformly positive, then there exists a positive eigenvector $c > 0$ and a unique “asymptotic growth coefficient” $\delta$, such that

$$\|fP_t - e^{\sigma m(f)}c\| \leq K^*e^{\sigma t}, \quad 0 \leq \sigma < \delta, \tag{20}$$

for every $f$, a suitable “effective initial size” $m(f)$, and $t \geq T$.

**Proof.** By Theorem 4, the discrete semigroup of $P^n_T = P_{nT}$ has the desired property; $m(f)$ is given by (17), with $e^{\delta T} = \gamma$. Furthermore, if $C$ is the “cone” of non-negative $f = f^+$ in $L$, then $CP_T$ has finite projective diameter $\Delta$. Hence, for any $t > nT$, we have

$$\Delta[CP_t] = \Delta[(CP_{t-nT})P_{nT}] \leq \Delta[CP_{nT}] \to 0 \tag{21}$$

where $\Delta[S] = \sup \{ \alpha \in \Theta(f, g; L^+) \}$, denotes the projective diameter of a cone $S$. It follows that the $c$ for $P_T$ (in the sense of Theorem 3) is the (unique) $c$ for $P_t$, which is also “uniformly positive,” and with the same $\delta$.

Finally, we can write $f = c + r$, where $r$ is in the complementary invariant subspace of $\S$. Applying (17), with $M = m(f)$, we get

$$\|fP^n_t - m(f)\gamma^n c\| \leq K^\rho^n c, \quad 0 \leq \rho < \gamma. \tag{22}$$

Hence, for any $t$ with $(n+1)T \leq t < (n+2)T$, we have

$$\|fP_t - e^{\delta t}m(f)c\| = \|(fP^n_t - m(f)\gamma^n c)P_{(t-nT)}\| \leq K^\rho^n\|c\|\|P_{t-nT}\|. \tag{23}$$

The uniform boundedness of $\|P_{t-nT}\|$ follows, however, from (16d). There follows

$$\|fP_t - e^{\delta t}m(f)c\| \leq K^*e^{\delta t}, \quad 0 \leq \sigma < \delta, \tag{24}$$

where $K^* = (K/\rho^2)\|c\|\sup_{T \leq t < 2T}\|P_t\|$, and $e^{\sigma T} = \rho$. This completes the proof.

Theorem 5 should be compared with the main result of Everett and Ulam(5). Our assumption of “uniform positivity” corresponds to uniform
mixing at a finite stage. For various physical applications, it would be desirable to weaken this hypothesis.

**Remark.** It is perhaps worth noting that all the preceding results apply to complete vector lattices\(^{(10)}\). The essential step is the following extension of Corollary 2 of Lemma 2, §6.

**Lemma 4.** Let \( L \) be any complete vector lattice. Relative to \( \theta(f, g; L^+) \), any connected component of \( L^+ \) is a complete metric space.

**Proof.** From any convergent subsequence \( \{g_k\} \), we can extract a hyper-convergent subsequence \( \{f_n\} = \{g_{k(n)}\} \), such that \( \theta(f_n, f_{n+1}; L^+) < 2^{-n} \). There follows, as in (13),

\[
|f_{n+1} - f_n| \leq ((R_n)^{1/2} - 1) |f_n| < 2^{-n}f_n, \quad f_n = |f_n| \in L^+.
\]

By the triangle inequality and induction, \( |f_{n+h} - f_n| \leq f_m/2^{n-1} \) for any \( m \leq n \); hence all \( f_n \) satisfy \( 0 < f_n < 4f_1 \). The vector lattice being complete, \( f_\infty = \bigvee f_n \) therefore exists, and \( |f_\infty - f_n| \leq |f_n|/2^{n-2} \leq f_1/2^{n-4} \), whence \( f_n \to f_\infty \) in the sense of relative uniform convergence and in the intrinsic topology.

**Harvard University,**

**Cambridge, Mass.**

\(^{(10)}\) In the sense of G. Birkhoff, *Lattice theory*, Chap. XV, §3.