PACIFIC JOURNAL OF MATHEMATICS
Vol. 116, No. 2, 1985

AN ANALOGUE OF LIAPOUNOFF'S CONVEXITY THEOREM FOR BIRNBAUM-ORLICZ SPACES AND THE EXTREME POINTS OF THEIR UNIT BALLS

I. GLICKSBERG

To Z. W. Birnbaum, on the occasion of his 80th birthday

Let (X, S, μ) be a non-atomic probability space. Our purpose is to note an analogue of Liapounoff's convexity theorem, as a statement about $L^\infty(\mu)$, for certain real Birnbaum-Orlicz spaces $L_\Phi(\mu)$, in particular reflexive ones, under the usual norms: the extreme points of the unit ball yield the full image of the ball under finite dimensional continuous linear maps.

1. Of course such a result is trivial if the ball is strictly convex (when each of its boundary points is extreme) since it just asserts that any finite codimensional closed subspace which meets the ball meets its extreme elements. But for $L_\Phi(\mu)$ the unit ball will have flat spots on its boundary corresponding to horizontal segments in the graph of $\varphi = \Phi'$, and the strong sort of density of extreme points the result implies is nontrivial. We shall obtain the analogue in fact as an application of Liapounoff's theorem (resulting from the use of support functionals suggested by [3]), and, although characterizations of the extreme points of the balls could be avoided, we shall obtain these too, so that Lindenstrauss' elegant proof of the Liapounoff result [5, 6] can also be applied.

Needless to say the assertion of the result makes sense for any Banach space, and fails if no extreme points exist; indeed it fails for $L^\infty(\mu)$ if our map is not w^* continuous. But it easily fails with that restriction, for example for the space of real measures on $[-2, -1] \cup [1, 2]$ and the map into $\mathbf{R}$ provided by $\nu \to \int \operatorname{sgn} t\, \nu(dt)$.

Finally we adapt the argument of [3] to one instance where neither of the approaches to the reflexive case applies (Theorem 2 below).

2. Let Φ and Ψ be dual Young's functions [2, 7] i.e., $\Phi(x) = \int_0^x \varphi(t)\,dt$, $\Psi(x) = \int_0^x \psi(t)\,dt$, where φ: $\mathbf{R}_+ \to \mathbf{R}_+$ is 0 at 0, non-decreasing and $\to \infty$ at ∞, while ψ is its "inverse", both taken continuous from the left for definiteness. As in [7] we complete the graph of φ by adding vertical segments at any discontinuities, and speak of the resulting curve as the extended graph of φ; then equality holds in Young's inequality

$$ab \leq \Phi(a) + \Psi(b), \quad \text{all } a, b \geq 0$$

iff (a, b) lies on the extended graph of φ. Assuming the convex function Φ has at most exponential growth at infinity (so $\Phi(2t) \leq M\Phi(t)$ for t large, the "Δ_2-condition" [**1, 2, 7**]) insures the class of real (or complex) functions

$$L_\Phi(\mu) = \left\{f: \int \Phi(|f|)\, d\mu < \infty\right\}$$

is a linear space; this becomes a Banach space when appropriately normed, in particular when

$$B_{\Phi,\mu} = \left\{f: \int \Phi(|f|)\, d\mu \leq 1\right\},$$

which is convex by Jensen's inequality, is taken as the closed unit ball (yielding the Luxemburg norm $\|\cdot\|$ [**2**]). The Δ_2 condition then also guarantees [**2**, p. 64] the dual is provided by the linear span $\mathbf{R}_+ L_\Psi(\mu)$ of $L_\Psi(\mu)$ under the pairing $\langle f, g\rangle = \int fg\, d\mu$. (Taking the polar $B^\circ_{\Psi,\mu}$ of $B_{\Psi,\mu}$ as the ball in $L_\Phi(\mu)$ gives an equivalent norm $\|\cdot\|_*$, the functional norm, and corresponding results hold then as well. It is probably worth noting that $L_\Phi(\mu) \subset L^1(\mu)$ always, since some line of positive slope lies under the graph of Φ and μ is finite.)

When both Φ and Ψ satisfy the Δ_2 condition $L_\Phi(\mu)$ and $L_\Psi(\mu)$ are thus a dual pair of reflexive Banach spaces (and conversely [**2**]), and we have the following, where S^e denotes the set of extreme points of S.

THEOREM 1. *If μ is a non-atomic probability measure and the dual Young's functions Φ, Ψ satisfy the Δ_2 condition, then for any continuous linear map ρ of $L_\Phi(\mu)$ to $\mathbf{R}^n$*

$$\rho\left(B^e_{\Phi,\mu}\right) = \rho\left(B_{\Phi,\mu}\right).$$

Finally, if $\{[a_j, b_j]\}$ are the maximal intervals of constancy[1] *of φ, and $b_j/a_j \geq \eta > 1$ for all j with $a_j > 0$, then*

$$\rho\left(B^{\circ e}_{\Psi,\mu}\right) = \rho\left(B^\circ_{\Psi,\mu}\right).$$

Assuming ρ is continuous is equivalent to assuming ρ weakly continuous, and $\rho(u) = \int u(t) y(t)\, \mu(dt)$, where $y = (y_1, \ldots, y_n)$ has $y_i \in L_\Psi(\mu)$.

[1]More precisely these correspond to maximal horizontal intervals in the extended graph (since $\varphi(a_j) < c_j = \varphi(a_j +)$ is possible).

In case the y_i are in fact bounded we need not assume reflexivity; only Φ need satisfy the Δ_2-condition.

THEOREM 2. *If μ is non-atomic and Φ satisfies the Δ_2-condition while ρ: $L_\Phi(\mu) \to \mathbf{R}^n$ is defined by bounded functions, then $\rho(B^e_{\Phi,\mu}) = \rho(B_{\Phi,\mu})$, and, under the additional hypothesis noted in Theorem* 1, $\rho(B^{\circ e}_{\Psi,\mu}) = \rho(B^{\circ}_{\Psi,\mu})$.

As we have remarked earlier the flat spots in the boundary of the balls needed to make the results non-trivial are provided by intervals of constancy for φ; these also figure in the following simple characterization of $B^e_{\Phi,\mu}$ which allows us easily to adapt Lindenstrauss' proof of Liapounoff's theorem [**6**, 5.5] to obtain the first half of Theorem 1. The corresponding and similar result for $B^{\circ e}_{\Psi,\mu}$, along with the (more complicated) proofs relating to $B^{\circ}_{\Psi,\mu}$, will be given in §6.

THEOREM 3. *Suppose μ is non-atomic and Φ satisfies the Δ_2-condition. Let $\{[a_j, b_j]\}$ be the maximal intervals of constancy of φ. Then $u_0 \in L_\Phi(\mu)$ is in $B^e_{\Phi,\mu}$ iff $\int \Phi(|u_0|)\, d\mu = 1$ and, a.e. on $|u_0|^{-1}[a_j, b_j]$, $|u_0(t)| = a_j$ or b_j, unless $a_j = 0$ when $|u_0(t)| = b_j$ instead.*

With this in hand one easily modifies the final step in Lindenstrauss' proof to obtain Theorem 1 (which applies since reflexivity provides the needed weak compactness): we have only to see an extreme point of the subset of $B_{\Phi,\mu}$ mapping onto a given point is extreme in $B_{\Phi,\mu}$. But if not by Theorem 3 $|u_0|^{-1}(a_j, b_j)$ has positive measure for some j, so $|u_0|^{-1}[a_j + \varepsilon, b_j - \varepsilon]$ also does for some $\varepsilon > 0$. Now one chooses a non-zero u supported by that set with $\operatorname{sgn} u = \operatorname{sgn} u_0$, $|u| \le \varepsilon$ and $\rho(u) = 0$, $\int \Phi(|u_0 \pm u|)\, d\mu = 1$, all of which are possible since Φ is linear on $[a_j, b_j]$ and $|u_0 + u| = (u_0 + u) \operatorname{sgn} u_0$, $|u_0 - u| = (u_0 - u)\operatorname{sgn} u_0$, while we can take $\int u \operatorname{sgn} u_0\, d\mu = 0$ and $\rho(u) = 0$ with u still non-trivial, indeed of modulus ε on $|u_0|^{-1}[a_j + \varepsilon, b_j - \varepsilon]$ by Liapounoff's theorem itself. Thus $\rho(u_0 \pm u) = \rho(u_0)$ and $u_0 \pm u \in B_{\Phi,\mu}$, our contradiction.

3. Theorem 3 easily follows from the next two lemmas (as we shall see later), which will prove basic to our argument.

LEMMA 1. *For any dual Young's functions Φ, Ψ and $0 \le v \in L_\Psi(\mu)$ with $\int \Psi(\lambda v)\, d\mu < \infty$ for $0 < \lambda < \lambda_1$, we have*

$$\int \Phi(\psi(\lambda v))\, d\mu < \infty \quad \text{for } 0 < \lambda < \lambda_1.$$

(In particular this applies for all $\lambda > 0$ if v is bounded or Ψ satisfies the Δ_2 condition, our applications in Theorems 2 and 1, respectively.)

For the proof, note that because

$$\int \Phi(\psi(\lambda v))\, d\mu + \int \Psi(\lambda v)\, d\mu = \int \lambda v \psi(\lambda v)\, d\mu$$

our hypothesis says the conclusion (for $\lambda < \lambda_1 - \varepsilon$) is equivalent to

$$\int v\psi(\lambda v)\, d\mu < \infty \quad (\text{for } 0 < \lambda < \lambda_1 - \varepsilon).$$

But since $\Psi(x) = \int_0^x \psi(t)\, dt$ is convex, as $h \searrow 0$

$$\frac{1}{h}\left[\Psi((\lambda + h)v(t)) - \Psi(\lambda v(t))\right] \searrow v(t)\psi(\lambda v(t) +)$$

while the difference quotient is bounded by $\frac{1}{\varepsilon}\left[\Psi((\lambda + \varepsilon)v(t)) - \Psi(\lambda v(t))\right]$ for $0 < h < \varepsilon$. So by dominated convergence

$$\int v\psi(\lambda v +)\, d\mu = \lim \frac{1}{h} \int \Psi((\lambda + h)v) - \Psi(\lambda v)\, d\mu$$

$$\leq \frac{1}{\varepsilon} \int \Psi((\lambda + \varepsilon)v) - \Psi(\lambda v)\, d\mu < \infty.$$

In essence we shall prove Theorem 1 by noting that for any $x_0 \in \rho(B_{\Phi,\mu})$ (or $\rho(B^{\circ}_{\Psi,\mu})$, later) there is an element u_0 of the boundary of $B_{\Phi,\mu}$ which maps onto x_0 (because ρ has a non-trivial kernel), whence x_0 lies in the image of a supporting subset of the boundary by Hahn-Banach. So it will suffice to see each supporting subset has the desired property for its extreme points. One is led thus to consider how an element of L_Ψ maximizes on $B_{\Phi,\mu}$, which we turn to next. We shall then give the proof of Theorems 1, 3, and finally Theorem 2, as they relate to $B_{\Phi,\mu}$.

Suppose a non-zero $v \in L_\Psi(\mu)$ is given and Ψ satisfies the Δ_2 condition or v is bounded. Then by Lemma 1 we have $\psi(\lambda|v|) \in L_\Phi(\mu)$ for all $\lambda > 0$. To see how

$$c = \max\left\{\int u(t)v(t)\,\mu(dt) : u \in B_{\Phi,\mu}\right\} \tag{1}$$

is achieved (if at all), note that if we could choose a $\lambda_0 > 0$ so that

$$\int \Phi\big(\psi(\lambda_0|v|)\big)\, d\mu = 1 \tag{2}$$

then for

$$u_0 = \psi(\lambda_0|v|)\,\operatorname{sgn} v$$

since

$$\lambda_0 v(t) u_0(t) = \Phi(|u_0(t)|) + \Psi(\lambda_0 |v(t)|) \quad \text{a.e.} \tag{3}$$

we have for all $u \in B_{\Phi,\mu}$

$$\begin{aligned} \int \lambda_0 v u_0 \, d\mu &= \int \Phi(|u_0|) \, d\mu + \int \Psi(\lambda_0|v|) \, d\mu = 1 + \int \Psi(\lambda_0 |v|) \, d\mu \\ &\geq \int \Phi(|u|) \, d\mu + \int \Psi(\lambda_0|v|) \, d\mu \geq \int \lambda_0 v u \, d\mu, \end{aligned} \tag{4}$$

so u_0 yields our maximum in (1). Moreover any maximizing u_1 must provide equality in (4) so it necessarily has

$$\int \Phi(|u_1|) \, d\mu = 1 \tag{5}$$

and also, because of equality in Young's inequality,

$$\begin{aligned} &(|u_1(t)|, \lambda_0 |v(t)|) \textit{ lies in the extended graph of } \varphi \textit{ and} \\ &\operatorname{sgn} u_1(t) = \operatorname{sgn} v(t) \text{ a.e. } (\textit{except where } uv = 0). \end{aligned} \tag{6}$$

Conversely (5) and (6) imply (3) and (4), hence that u_1 maximizes.

In fact, precisely because of the horizontal segments in the graph of φ, we may not have a λ_0 yielding (2). Setting

$$b(\lambda) = \int \Phi(\psi(\lambda|v|)) \, d\mu$$

and noting that $b(\lambda) \nearrow \infty$ as $\lambda \nearrow \infty$ since $|v| > 0$ on a set of positive measure, clearly $\lambda_0 = \sup\{\lambda > 0 : b(\lambda) \leq 1\}$ yields our candidate; either $b(\lambda_0) = 1$ or $b(\lambda_0) < 1 \leq b(\lambda_0 +) = \lim_{\lambda \searrow \lambda_0} b(\lambda)$. By monotone convergence and our assumption that $\psi(t-) = \psi(t)$

$$b(\lambda_0 -) = \int \Phi(\psi(\lambda_0 |v|)) \, d\mu = b(\lambda_0);$$

because of Lemma 1, by dominated convergence

$$b(\lambda_0 +) = \int \Phi(\psi(\lambda_0 |v(t)| +)) \, \mu(dt).$$

Let $D = \{c_j\}$ be the set of discontinuities of ψ, $a_j = \psi(c_j -) = \psi(c_j)$, $b_j = \psi(c_j +)$ (so (a_j, c_j) and (b_j, c_j) are the endpoints of the corresponding horizontal segments in the graph of φ). Then, with $D' = (\lambda_0|v|)^{-1}(D)$,

$$b(\lambda_0 +) = \int_{X \setminus D'} \Phi(\psi(\lambda_0 |v|)) \, d\mu + \sum_j \int_{(\lambda_0|v|)^{-1}(c_j)} \Phi(b_j) \, d\mu \tag{7}$$

while

$$b(\lambda_0) = \int_{X \setminus D'} \Phi(\psi(\lambda_0|v|))\, d\mu + \sum_j \int_{(\lambda_0|v|)^{-1}(c_j)} \Phi(a_j)\, d\mu.$$

Evidently since $b(\lambda_0 +) \geq 1$ we can now alter $u_0 = \psi(\lambda_0|v|) \operatorname{sgn} v$ on D' so as to obtain both (5) and (6), and thus (3) and (4) for the resulting u. Indeed if $b(\lambda_0 +) = 1$ we simply increase $|u_0|$ to b_j on $(\lambda_0|v|)^{-1}(c_j)$ for each j; if not we choose the largest N for which the sum of the first N terms on the right side of (7) is < 1, increase $|u_0|$ to b_j on $(\lambda_0|v|)^{-1}(c_j)$, $j \leq N$, while, since μ is non-atomic, on an appropriate subset of $(\lambda_0|v|)^{-1}(c_{N+1})$ we increase $|u_0|$ to b_{N+1}.

We can summarize most of the preceding as

LEMMA 2. *Suppose v is a non-zero element of $L_\Psi(\mu)$, and v is bounded, or Ψ satisfies the Δ_2 condition. Let*

$$\lambda_0 = \sup\left\{\lambda > 0 : \int \Phi(\psi(\lambda|v|))\, d\mu \leq 1\right\}.$$

Then there are u in $B_{\Phi,\mu}$ which provide the maximum in (1), *and they are just those u satisfying* (5) *and* (6). *Moreover,*

$$\|v\|_* = \frac{1}{\lambda_0}\left(1 + \int \Psi(\lambda_0|v|)\, d\mu\right) = \min_{\lambda > 0} \frac{1}{\lambda}\left(1 + \int \Psi(\lambda|v|)\, d\mu\right).$$

For the final assertion, note that by (4) we have

$$\|v\|_* = (1/\lambda_0)(1 + \textstyle\int \psi(\lambda_0|v|)\, d\mu),$$

and we only have to see $h(\lambda) = (1/\lambda)(1 + \int \Psi(\lambda|v|)\, d\mu)$ minimizes at $\lambda = \lambda_0$. But h has a left hand derivative everywhere which is

$$\begin{aligned} -\frac{1}{\lambda^2}&\left(1 + \int \Psi(\lambda|v|)\, d\mu\right) + \frac{1}{\lambda}\int (|v|\psi(\lambda|v|)\, d\mu) \\ &= \frac{1}{\lambda^2}\left(\int \lambda|v|\psi(\lambda|v|)\, d\mu - 1 - \int \Psi(\lambda|v|)\, d\mu\right) \\ &= \frac{1}{\lambda^2}\left(\int \Phi(\psi(\lambda|v|))\, d\mu - 1\right) \end{aligned}$$

(using equality in Young's inequality), so the result is ≤ 0 for $\lambda < \lambda_0$, > 0 for $\lambda > \lambda_0$. Since the convex function $\lambda \to \int \Psi(\lambda|v|)\, d\mu$ is absolutely continuous, so is h. Thus, as the integral of this derivative, it has its minimum at λ_0.

REMARK. For later use we should note that λ_0 alone provides our minimum unless our derivative is zero on an interval $(\lambda_0 - \varepsilon, \lambda_0]$ where the range of $\lambda|v|$ lies entirely in the intervals of constancy of ψ.

We can now prove Theorem 3. First suppose $\int \Phi(|u_0|)\, d\mu = 1$ and $|u_0(t)| = a_j$ or b_j a.e. on $|u_0|^{-1}[a_j, b_j]$ (unless $a_j = 0$ when $|u_0(t)| = b_j$). Then for $v = \varphi(|u_0|) \operatorname{sgn} u_0$, which lies in $L_\Psi(\mu)$ by Lemma 1 (with Φ and Ψ interchanged) we have

$$u_0 v = \Phi(|u_0|) + \Psi(|v|) \quad \text{a.e.}$$

so that

$$\int u_0 v\, d\mu = \int \Phi(|u_0|)\, d\mu + \int \Psi(|v|)\, d\mu = 1 + \int \Psi(|v|)\, d\mu \geq \int uv\, d\mu$$

for any $u \in B_{\Phi,\mu}$. Thus if $u_0 = \frac{1}{2}(u_1 + u_2)$ with $u_i \in B_{\Phi,\mu}$, $i = 1, 2$, each u_i must maximize the right-most term, so again provide equality a.e. in Young's inequality: $(|u_i(t)|, |v(t)|)$ lies on the extended graph of φ and $\operatorname{sgn} u_i(t) = \operatorname{sgn} v(t) = \operatorname{sgn} u_0(t)$ a.e. except where $u_i v = 0$. Evidently then $u_i(t) = u_0(t)$ a.e. except on the sets $|u_0|^{-1}[a_j, b_j]$ where the extremity of values of $|u_0|$ forces the same extremity for $|u_1|$, $|u_2|$. So $u_1 = u_0$, and u_0 is extreme.

Conversely suppose $u \in B^e_{\Phi,\mu}$. Then $\int \Phi(|u|)\, d\mu < 1$ implies $\int \Phi(\lambda|u|)\, d\mu < 1$ for some $\lambda > 1$ and u is interior to a segment in $B_{\Phi,\mu}$. So $\int \Phi(|u|)\, d\mu = 1$. Again, if for some j, $|u|$ fails to assume the values a_j and b_j on a set of positive measure in $|u|^{-1}[a_j, b_j]$ then $\mu(|u|^{-1}[a_j + \varepsilon, b_j - \varepsilon]) > 0$ for some $\varepsilon > 0$. Now we partition $F = |u|^{-1}[a_j + \varepsilon, b_j - \varepsilon]$ into disjoint subsets F_1 and F_2 of equal measure, and, since Φ is linear on $[a_j, b_j]$, for $|\theta| \leq \varepsilon$ we have

$$\begin{aligned}
&\int \Phi\big(|u + \theta(\chi_{F_1} - \chi_{F_2}) \operatorname{sgn} u|\big)\, d\mu \\
&\qquad = \int_{X \setminus F} \Phi(|u|)\, d\mu + \int_{F_1} \Phi(|u| + \theta)\, d\mu + \int_{F_2} \Phi(|u| - \theta)\, d\mu \\
&\qquad = \int \Phi(|u|)\, d\mu = 1.
\end{aligned}$$

Hence for $h = (\chi_{F_1} - \chi_{F_2}) \operatorname{sgn} u$ we have $u \pm \varepsilon h \in B_{\Phi,\mu}$ so u cannot be extreme, completing our proof.

4. In effect we have also proved Theorem 1 if we adapt Lindenstrauss' proof as indicated before. On the other hand, because ρ has a nontrivial kernel, as noted earlier any $x_0 \in \rho(B_{\Phi,\mu})$ is the image of an element u_0 of the boundary of $B_{\Phi,\mu}$. Thus we can invoke the Hahn-Banach theorem to produce a functional, given by some $v \in L_\Phi(\mu)$ (by the Δ_2-condition for Φ), which maximizes over $B_{\Phi,\mu}$ at u_0. Since Ψ also satisfies the Δ_2 condition, Lemma 2 implies the support set v provides consists of those u in $B_{\Phi,\mu}$ satisfying (5) and (6). The fact that an extreme point lies therein which also maps onto x_0 is in fact a consequence of Liapounoff's theorem itself. For from our discussion of the maximization question (1), we know all such $u = u_0$ a.e. off $D' = (\lambda_0|v|)^{-1}(D)$; the fact that $\int \Phi(|u|)\, d\mu = 1 = \int \Phi(|u_0|)\, d\mu$ says

$$\int_{D'} \Phi(|u|)\, d\mu = \int_{D'} \Phi\left(|u_0|\right) d\mu.$$

Since $\operatorname{sgn} u = \operatorname{sgn} u_0 = \operatorname{sgn} v$, and Φ is linear on $[a_j, b_j]$, this is equivalent to the fact that, for certain $g, h \in L^1(\mu)$,

$$\int_{D'} \Phi\left(|u_0|\right) d\mu = \int_{D'} g|u| + h\, d\mu = \int_{D'} (ug \operatorname{sgn} v + h)\, d\mu.$$

Thus our supporting u, determined off D', must satisfy two constraints on D':

(9) a.e. on $(\lambda_0|v|)^{-1}(c_j)$, $a_j \le |u| \le b_j$ and $\operatorname{sgn} u = \operatorname{sgn} v$ (unless $uv = 0$), or equivalently, unless $c_j = 0$, $a_j \le u \operatorname{sgn} v \le b_j$.

$$\int_{D'} ug \operatorname{sgn} v\, d\mu = k. \tag{10}$$

In order to map onto x_0 any such u would also satisfy

$$\int_{D'} uy\, d\mu = x_0 - \int_{X\setminus D'} u_0 y\, d\mu, \tag{11}$$

where $y = (y_1, \ldots, y_n) \in (L_\Psi(\mu))^n$ gives rise to ρ.

But now the variation possible on D' is trivially affinely equivalent to that in an $(n + 1)$-dimensional Liapounoff problem: setting

$$\tilde{u} = \frac{1}{b_j - a_j}\left(u \operatorname{sgn} v - \frac{1}{2}(a_j + b_j)\right) \quad \text{on } \left(\lambda_0|v|\right)^{-1}(c_j) \text{ if } c_j \neq 0$$

or $\tilde{u} = (1/b_j)u \operatorname{sgn} v$ on that set if $c_j = 0$, we are concerned precisely with an $n + 1$ dimensional image of $\tilde{u}$ in ball $L^\infty(\mu_{D'})$. Hence Liapounoff's theorem applies to assert (10) and (11) can be satisfied on D' by u with $u \operatorname{sgn} v$ assuming the extreme values a_j or b_j a.e. on $(\lambda_0|v|)^{-1}(c_j)$ (except when $c_j = 0$ and then $u = \pm b_j$ clearly). Extending by u_0 off D' we now have a u mapping onto x_0 which is extreme in $B_{\Phi,\mu}$ by our old observation: $u = \frac{1}{2}(u_1 + u_2)$ for $u_i \in B_{\Phi,\mu}$ implies $u_i = u_0 = u$ a.e. off D' since u_i lies in our support set where v maximizes, while the extremity of u's values on $(\lambda_0|v|)^{-1}[a_j, b_j]$ forces $u_i(t) = u(t)$ a.e. there too.

5. In the setting of Theorem 2 we cannot appeal wholly to either of the approaches used for Theorem 1: without weak compactness we cannot invoke the Krein-Milman Theorem as in Lindenstrauss' proof, and the approach through Hahn-Banach and maximization cannot use the boundedness of our components y_i. What will exploit that is the argument of [3] giving a controlled approach to Liapounoff's result: given x_0 in $\rho(B_{\Phi,\mu})$ we choose a minimal set $E \in S$ capturing x_0 in the sense that x_0 lies in

$$K_y(E) = \left\{\rho(u\chi_E) = \int_E uy\,d\mu : u \in B_{\Phi,\mu}\right\}; \tag{12}$$

then x_0 must lie in the boundary of this set so that an element θ of $\mathbf{R}^n$ provides a supporting functional there, and $\theta \cdot y \in L_\Psi(\mu)$ then provides a bounded function against which any $u\chi_E$ mapping onto x_0 must maximize. We then shall see there is a $u\chi_E \in B^e_{\Phi,\mu_E}$ mapping onto x_0, and we extend this to all of X by an appropriate detour. The details will require some technical refinements of the arguments of [3], all possible precisely because of the assumed boundedness of y.

Let $K^e_y(E)$ denote the subset of $K_y(E)$ defined similarly with $B^e_{\Phi,\mu}$ in place of $B_{\Phi,\mu}$. In order to proceed we first need to see $K_y(E)$ (trivially convex) *is always closed.* We can assume $(y_1\chi_E, \ldots, y_n\chi_E)$ are linearly independent (or restrict our attention to a lower dimensional map). But any point x_0 in $\partial K_y(E)$ lies in a support set of $K_y(E)^-$ given by a unit vector θ in $\mathbf{R}^n$, so $x_0 = \rho(u\chi_E)$ for some u only if $u\chi_E$ provides the supremum of $\int u\chi_E\theta \cdot y\,d\mu$ over $B_{\Phi,\mu}$, and such u exist by Lemma 2. Moreover our maximizing u has, for some $\lambda_0 = \lambda_\theta > 0$, $(|u(t)|, \lambda_\theta|\theta \cdot y(t)|)$ in the extended graph of φ a.e., and so is bounded since y is.

In fact there is a bound M independent of our unit vector $\theta \in S^{n-1}$: for $\mu\{t \in E: |\theta \cdot y(t)| > 0\} > 0$ by independence, so given θ,

$$\mu\{t \in E: |\theta \cdot y(t)| > 2\varepsilon\} > \delta > 0$$

for some ε and $\delta > 0$, and replacing 2ε by ε the inequality holds for nearby θ in S^{n-1}. Compactness of S^{n-1} now shows we have $\mu\{t \in E: |\theta \cdot y(t)| > \varepsilon\} > \delta > 0$ for some ε, $\delta > 0$ and every θ. But now

$$1 \geq \int \Phi(\psi(\lambda_\theta|\theta \cdot y|))\,d\mu \geq \Phi(\psi(\lambda_\theta\varepsilon))\delta$$

shows the existence of an upper bound m for λ_θ, hence an upper bound $M = \psi(mM_1 +)$ for $|u| \leq \psi(\lambda_\theta|\theta \cdot y| +)$ if M_1 bounds $\|y\|$.

Because of this no $u \in \chi_E \cdot B_{\Phi,\mu}$ which assumes values of modulus $> M$ on a set of positive measure can have $\rho(\chi_E u) = \rho(u) \in \partial K_y(E)$ since then it would maximize for some θ. Consequently $\partial K_y(E)$ remains unchanged if we alter φ on $(2M, \infty)$ so as to obtain Φ_1, Ψ_1 both satisfying the Δ_2-condition, as we can trivially do; in fact we can take φ_1 linear on $(2M, \infty)$ and thus insure that $\lim_{t\to\infty} \Phi_1(t)/t = \infty$. If we now set

$$K_y^1(E) = \left\{ \int_E uy\, d\mu : u \in B_{\Phi_1,\mu} \right\}$$

then this is a compact convex set because of the weak compactness of $B_{\Phi_1,\mu}$ reflexivity insures, and by the preceding all of $\partial K_y^1(E)$ is provided by u's bounded by M so necessarily lies in $K_y(E)$. We conclude $K_y(E) = K_y^1(E)$, and $K_y(E)$ is closed as asserted. (Note that our M and Φ_1 obtained for E can be used equally well for its subsets to obtain the same conclusions.)

Now because of this and the boundedness of y we can also see $K_y(\bigcap E_n) = \bigcap K_y(E_n)$ for a decreasing sequence $\{E_n\}$, as we must to construct our minimal E. To obtain this we only have to show $x_0 \in \bigcap K_y(E_n)$ lies in $K_y(E)$, $E = \bigcap E_n$. But having chosen M and Φ_1 corresponding to E_1 as above, we have $x_0 = \rho(u_n \chi_{E_n})$ for each n, with $u_n \in B_{\Phi_1,\mu}$ since $K_y(E_n) = K_y^1(E_n)$; since [2, p. 94] $\lim_{t\to\infty} \Phi_1(t)/t = \infty$ implies $\{u_n\}$ is equi-integrable $\|\rho(u_n\chi_{E_n}) - \rho(u_n\chi_E)\| = \|\int_{E_n\setminus E} u_n y\, d\mu\| \to 0$ since y is bounded. Thus $x_0 = \lim \rho(u_n\chi_E)$ lies in $K_y(E)$ since the set is closed and $\rho(u_n\chi_E) \in K_y^1(E) = K_y(E)$.

We can now see how the argument of [3] adapts. Suppose $x_0 \in K_y(E)$. We let

$$c_1 = \inf\{\mu(E) : E \subset E_1 = X, x_0 \in K_y(E)\}$$

and choose an E_2 from the competing E so that $c_1 \le \mu(E_2) \le c_1 + \varepsilon^{-1}(\mu(E_1) - c_1)$. Continuing we obtain a decreasing sequence $\{E_k\}$ and a non-decreasing sequence $\{c_k\}$ with

$$c_k = \inf\{\mu(E) : E \subset E_k, x_0 \in K_y(E)\} \le \mu(E_{k+1})$$
$$\le c_k + 2^{-k}(\mu(E_k) - c_k).$$

For $E = \bigcap E_k$ we know by the preceding that $K_y(E) = \bigcap K_y(E_k)$, so $x_0 \in K_y(E)$. As a consequence

$$x_0 \in \partial K_y(E) \tag{13}$$

since otherwise x_0 is interior to $K_y(E)$ and we can remove a bit of E to obtain a subset F of smaller measure with $x_0 \in K_y(F)$; then $\mu(F) < c_k$ for some k since $\mu(F) < \mu(E) \le \lim c_k$, contradicting the definition of c_k.

Because of (13) we have a $\theta \in \mathbf{R}^n$ supporting $K_y(E)$ at x_0, and by (12)

$$\theta \cdot x_0 = \max\left\{\int_E u(t)\theta \cdot y(t)\, \mu(dt) : u \in B_{\Phi,\mu}\right\}. \tag{14}$$

Now exactly the argument of §4, with μ replaced by μ_E and v by $\theta \cdot y$, shows that

$$x_0 = \int_E u(t) y(t)\, \mu(dt)$$

where $u \in B^e_{\Phi,\mu_E}$; indeed $u \operatorname{sgn} \theta \cdot y$ assumes extreme values wherever any variation is possible for any μ maximizing (14). If $u(E) = 1$ then our u is in $B^e_{\Phi,\mu}$ since $u = \frac{1}{2}(u_1 + u_2)$ implies $u_1 = u_2 = u$ a.e. on E. But if $\mu(E) < 1$ the fact that $\int_E \Phi(|u|)\, d\mu = 1$ insures that the only extensions of $u|_E$ to all of X which lie in $B_{\Phi,\mu}$ must have $\int_{X \setminus E} \Phi(|u|)\, d\mu = 0$, so u (and u_1, u_2) must vanish a.e. on $X \setminus E$ unless $0 \in D$ and we have an initial interval of constancy $[0, b]$ of φ, $b > 0$.

But now we can choose an extension of u with

$$\int_{X \setminus E} uy\, d\mu = 0, \tag{15}$$

hence $\rho(u) = x_0$, and $|u| = b$ a.e. on $X \setminus E$ (so clearly extreme in $B_{\Phi,\mu}$) by exactly the final part of the argument of [3] establishing the version [3, (4)] of (14) for u in the unit ball of $L^\infty(\mu_{X \setminus E})$. Our proof of Theorem 2 for $B_{\Phi,\mu}$ is now complete.

It might be noted that the role of the induction step in our proof [3] has been taken by Liapounoff's Theorem itself.

6. We now turn to $B^\circ_{\Phi,\mu}$ and the functional norm $\|\ \|_*$ on $L_\Phi(\mu)$. In order to obtain analogous results we consider the problem of maximizing a functional given by $v \in L_\Psi(\mu)$ over $B^\circ_{\Psi,\mu}$. Recall that by Lemma 2 (with Φ and Ψ interchanged) we have

$$\begin{aligned} \|u\|_* &= \frac{1}{\lambda_u}\left(1 + \int \Phi(\lambda_u |u|)\, d\mu\right) \\ &\text{where } \lambda_u = \sup\left\{\lambda > 0 : \int \Psi(\varphi(\lambda |u|))\, d\mu\right\}, \end{aligned} \tag{16}$$

if Φ satisfies the Δ_2 condition or u is bounded.

LEMMA 2°. *Suppose Φ and Ψ satisfy the Δ_2-condition or $v \in L_\Psi(\mu)$ is bounded, while $\int \Psi(|v|)\, d\mu = 1$ (so $\|v\| = 1$). Then there are u in $B^\circ_{\Psi,\mu}$ which* *provide*

$$\max\left\{\int uv\,d\mu\colon u \in B^{\circ}_{\Psi,\mu}\right\}$$

and they are precisely those with $\|u\|_* = 1$ *and, a.e.,* $(\lambda_u|u(t)|, |v(t)|)$ *in the extended graph of* φ *and* $\arg u(t) = \arg v(t)$ *(except where* $uv = 0$*). Finally* $u \to \lambda_u^{-1}$ *is affine on our support set except when* $|v|$ *has its range entirely in the set of discontinuities* $D = \{c_j\}$ *of* ψ*, i.e.* $\mu(|v|^{-1}(D)) = 1$.

Let $\lambda_0 = 1 + \int \Phi(\psi(|v|))\,d\mu$, which is finite by Lemma 1 (using our hypothesis on Ψ and v), and let $u_0 = (1/\lambda_0)\psi(|v|)\operatorname{sgn} v$, which is bounded iff v is. Then since $\lambda_0|u_0| = \psi(|v|)$

$$\int \lambda_0 u_0 v\,d\mu = \int \Phi(\lambda_0|u_0|)\,d\mu + \int \Psi(|v|)\,d\mu = \int \Phi(\lambda_0|u_0|)\,d\mu + 1 = \lambda_0,$$

so

$$\int u_0 v\,d\mu = 1. \tag{17}$$

We claim $\lambda_0 = \lambda_{u_0}$ (as in (16)). Indeed $|v| = \varphi(\lambda_0|u_0|)$ a.e. except where ψ^{-1} is not well defined; if $D' = \{c_j'\}$ are the discontinuities of φ, and (c_j', a_j'), (c_j', b_j') the endpoints of the corresponding vertical segments in the extended graph of φ, then on $|v|^{-1}[a_j', b_j']$, $\lambda_0|u_0| = c_j'$ so $\varphi(\lambda_0|u_0|) = \varphi(c_j') = a_j' \le |v| \le b_j'$. Thus

$$\int \Psi(\varphi(\lambda_0|u_0|))\,d\mu \le \int \Psi(|v|)\,d\mu = 1. \tag{18}$$

On the other hand no $\lambda > \lambda_0$ yields this, for as in Lemma 2

$$\lim_{\lambda \searrow \lambda_0} \int \Psi(\varphi(\lambda|u_0|))\,d\mu = \int_{X\setminus(\lambda_0|u_0|)^{-1}(D')} + \sum_j \int_{(\lambda_0|u_0|)^{-1}(c_j')} \Psi(b_j')\,d\mu$$

$$\ge \int \Psi(|v|)\,d\mu = 1.$$

So $\lambda_0 = \lambda_{u_0}$.

As a consequence, since $\lambda_0|u_0| = \psi(|v|)$, by (16) and the definition of λ_0

$$\|u_0\|_* = \frac{1}{\lambda_0}\left(1 + \int \Phi(\lambda_0|u_0|)\,d\mu\right) = \frac{1}{\lambda_0}\left(1 + \int \Phi(\psi(|v|))\,d\mu\right) = 1$$

so $u_0 \in B^{\circ}_{\Psi,\mu}$. By (17) now $\int u_0 v\,d\mu = 1 = \|u_0\|_*$, and since $\int uv\,d\mu \le \|u\|_*\|v\| = \|u\|_*$, for any $u \in B^{\circ}_{\Psi,\mu}$ we have

$$\int uv\,d\mu \le 1 = \int u_0 v\,d\mu$$

and u_0 provides our maximum. Moreover any maximizing u in $B^{\circ}_{\Psi,\mu}$ now clearly has $\int uv\,d\mu = 1 = \|u\|_*$ so $\lambda_u = 1 + \int \Phi(\lambda_u|u|)\,d\mu$ by (16). Consequently

$$\text{(19)} \qquad \int \lambda_u uv\,d\mu = \lambda_u = 1 + \int \Phi(\lambda_u|u|)\,d\mu$$

$$= \int \Psi(|v|)\,d\mu = \int \Phi(\lambda_u|u|)\,d\mu$$

and again by equality in Young's inequality we see u must satisfy our necessary conditions.

Conversely these imply (19), hence $\int uv\,d\mu = 1 = \|u\|_*$, whence u maximizes.

For the final assertion, since $\Phi' = \varphi = c_j$ on $\varphi^{-1}(c_j) = [a_j, b_j]$, and $\lambda_u|u| = \psi(|u|)$ on $X \setminus |v|^{-1}(D)$ where $D = \{c_j\}$, we have

$$\text{(20)} \qquad 1 = \frac{1}{\lambda_u}\left(1 + \int \Phi(\lambda_u|u|)\,d\mu\right)$$

$$= \frac{1}{\lambda_u}\left[1 + \int_{X\setminus|v|^{-1}(D)} \Phi(\psi|v|)\,d\mu + \sum_j \int_{|v|^{-1}(c_j)} \Phi(a_j) + c_j(\lambda_u|u| - a_j)\,d\mu\right]$$

$$= \frac{K}{\lambda_u} + \sum_j \int_{|v|^{-1}(c_j)} c_j|u|d\mu = \frac{K}{\lambda_u} + \sum_j \int_{|v|^{-1}(c_j)} c_j u \operatorname{sgn} v\,d\mu$$

so that $u \to 1/\lambda_u$ is affine on our support set except where $K = 0$. But since $\Phi(a_j) + \Psi(c_j) = a_j c_j$ (because (c_j, a_j) is on the extended graph of φ)

$$K = 1 + \int_{X\setminus|v|^{-1}(D)} \Phi(\psi(|v|))\,d\mu + \sum_j \int_{|v|^{-1}(c_j)} \Phi(a_j) - c_j a_j\,d\mu$$

$$= 1 + \int_{X\setminus|v|^{-1}(D)} \Phi(\psi|v|)\,d\mu + \sum_j \int_{|v|^{-1}(c_j)} \Psi(c_j)\,d\mu$$

$$= 1 + \int_{X\setminus|v|^{-1}(D)} \Phi(\psi|v|)\,d\mu - \int_{|v|^{-1}(D)} \Psi(|v|)\,d\mu.$$

Since $\int \Psi(|v|)\,d\mu = 1$ we conclude $u \to 1/\lambda_u$ is affine on our support set except precisely when $\mu(|v|^{-1}(D)) = 1$. Finally, as we shall need later $K = 0$ is also equivalent by (20) to

$$(21)\qquad 1 = \sum_j \int_{|v|^{-1}(c_j)} c_j|u|\,d\mu = \int_{|v|^{-1}(D)} uv\,d\mu$$

since $\operatorname{sgn} u = \operatorname{sgn} v$.

We can now almost characterize $B^{\circ e}_{\Psi,\mu}$, but the result is incomplete and certainly more complicated than its predecessor.

THEOREM 3°. *Suppose Φ and Ψ satisfy the Δ_2-condition and $\{[a_j, b_j]\}$ are the maximal intervals of constancy of φ. Let $u \in L_\Phi(\mu)$ and $\lambda_u = \sup\{\lambda > 0\colon \int \Psi(\varphi(\lambda|u|))\,d\mu \le 1\}$. Then $u \in B^{\circ e}_{\Psi,\mu}$ if* (i) $\|u\|_* = 1$, (ii) *a.e. on $|u|^{-1}[\lambda_u^{-1}a_j, \lambda_u^{-1}b_j]$, $|u(t)| = \lambda_u^{-1}a_j$ or $\lambda_u^{-1}b_j$ (unless $a_j = 0$ when $|u(t)| = \lambda_u^{-1}b_j$ instead), and* (iii) *if $\mu(|u|^{-1}(\bigcup[\lambda_u^{-1}a_j, \lambda_u^{-1}b_j])) = 1$ and $\int \Psi(\varphi(\lambda_u|u|))\,d\mu = 1$ then $\mu(|u|^{-1}(\lambda_u^{-1}a_j)) > 0$ for some j as well. Conversely $u \in B^{\circ e}_{\Psi,\mu}$ implies* (i) *and* (ii), *and, provided $b_j/a_j \ge \eta > 1$ for all j with $a_j > 0$, it also implies* (iii).

In fact if u is bounded we do not need the Δ_2-conditions. Note that in particular if there are only finitely many intervals of constancy we have a complete characterization.

Suppose first our u satisfies (i)–(iii). Exactly as in the proof of Lemma 2 (with Φ and Ψ interchanged) we can alter $v = \varphi(\lambda_u|u|)\operatorname{sgn} u$ so as to obtain $v \in L_\Psi(\mu)$ with $\int \Psi(|v|)\,d\mu = 1$ while $(\lambda_u|u(t)|, |v(t)|)$ lies on the extended graph of φ and $\operatorname{sgn} v = \operatorname{sgn} u$ a.e.. Note moreover that (with c_j', a_j', b_j' as in Lemma 2°) in altering v on the sets $(\lambda_u|u|)^{-1}(c_j')$, unless we are forced to increase $|v|$ to b_j' for all j (just the case where

$$\left.\int \Psi(\varphi(\lambda_u|u(t)| +))\,d\mu = 1\right)$$

we can always alter $|v|$ so that on some such set $a_j' < |v| < b_j'$ on a set of positive measure, and thus have $\mu(|v|^{-1}(D)) < 1$ necessarily (since such values cannot give rise to points on the horizontal segments of the extended graph of φ). So we have

$$(22)\qquad \mu\big(|v|^{-1}(D)\big) < 1, \quad\text{or}\quad \int \Psi\big(\varphi\big(\lambda_u|u(t)| +\big)\big)\,\mu(dt) = 1.$$

Now by Lemma 2°, as a functional v maximizes over $B^{\circ}_{\Psi,\mu}$, at our u, hence if $u = \frac{1}{2}u_1 + \frac{1}{2}u_2$ with $u_i \in B^{\circ}_{\Psi,\mu}$, necessarily at each u_i as well. So $(\lambda_{u_i}|u_i(t)|, |v(t)|)$ lies on the extended graph and $\operatorname{sgn} u_i = \operatorname{sgn} v = \operatorname{sgn} u$ a.e., and $\lambda_{u_0}u_i = \lambda_u u$ except on the $|v|^{-1}(c_j)$. Hence $\lambda_{u_i}u_i = \lambda_u u$ except on the $|v|^{-1}(c_j)$ where $a_j \le \lambda_{u_i}|u_i| \le b_j$, or $\lambda_{u_i}^{-1}a_j \le |u_i| \le \lambda_{u_i}^{-1}b_j$.

We now have two cases: either (21) fails ($K \neq 0$) and $u \to \lambda_u^{-1}$ is affine, or (21) holds. In the first case since $\lambda_u^{-1} = \frac{1}{2}(\lambda_{u_1}^{-1} + \lambda_{u_2}^{-1})$ we have

$$\tfrac{1}{2}\lambda_{u_1}^{-1}a_j + \tfrac{1}{2}\lambda_{u_2}^{-1}a_j = \lambda_u^{-1}a_j \leq \tfrac{1}{2}|u_1| + \tfrac{1}{2}|u_2| = |u| \leq \lambda_u^{-1}b_j$$

on $|v|^{-1}(c_j)$ and thus from the assumed extremity of the values of $|u|$ on that set we must have $|u_1| = \lambda_{u_1}^{-1}a_j$ and $|u_2| = \lambda_{u_2}^{-1}a_j$ where $|u| = \lambda_u^{-1}a_j$ (and similarly for the b_j), which says $\lambda_{u_1}|u_1| = \lambda_u|u| = \lambda_{u_2}|u_2|$ a.e. on all X. But now $\lambda_{u_i} = 1 + \int \Phi(\lambda_{u_i}|u_i|)\,d\mu = 1 + \int \Phi(\lambda_u|u|)\,d\mu = \lambda_u$ and we conclude $u_i = u$, so u is extreme, as desired.

In the second case, where by (21) $\int_{|v|^{-1}(D)} uv\,d\mu = 1$ and also $\mu(|v|^{-1}(D)) = 1$, we necessarily have $v = v\chi_{|v|^{-1}(D)}$ so $\int \Psi(|v\chi_{|v|^{-1}(D)}|)\,d\mu = 1$. Since $|v| = \varphi(\lambda_u|u|)$ on $|v|^{-1}(c_j)$ except when (a_j is a discontinuity of φ and)

$$\mu\left(|u|^{-1}\left(\lambda_u^{-1}a_j\right)\right) > 0, \tag{23}$$

when (23) fails for all j we obtain $\int \Psi(\varphi(\lambda_u|u|))\,d\mu = 1$. Thus by (iii) (23) holds for some j.

On the other hand when $\mu(|v|^{-1}(D)) = 1$ we cannot have $\lambda_u|u| \leq \eta b_j$ for all j, with $\eta < 1$, since then $\lambda = \eta^{-1}\lambda_u$ has the property that $a_j < \lambda|u| \leq b_j$ on $|v|^{-1}(c_j)$, whence $\varphi(\lambda|u|) = \varphi(a_j +) = c_j$, all j, so $\int \Psi(\varphi(\lambda|u|))\,d\mu = \int \Psi(\varphi(\lambda_u|u| +))\,d\mu = 1$ by (22), which contradicts the definition of λ_u. Consequently for $\varepsilon > 0$ we have $\lambda_u|u| > (1 - \varepsilon)b_j$ on a set of positive measure for some j, and thus

$$\lambda_u^{-1}(1 - \varepsilon)b_j < |u| = \tfrac{1}{2}|u_1| + \tfrac{1}{2}|u_2| \leq \tfrac{1}{2}\lambda_{u_1}^{-1}b_j + \tfrac{1}{2}\lambda_{u_2}^{-1}b_j$$

on that set. So $\lambda_u^{-1}(1 - \varepsilon) \leq \frac{1}{2}\lambda_{u_1}^{-1} + \frac{1}{2}\lambda_{u_2}^{-1}$ for any $\varepsilon > 0$, and $\lambda_u^{-1} \leq \frac{1}{2}\lambda_{u_1}^{-1} + \frac{1}{2}\lambda_{u_2}^{-1}$.

By (23) we have the reverse inequality: for

$$\tfrac{1}{2}\lambda_{u_1}^{-1}a_j + \tfrac{1}{2}\lambda_{u_2}^{-1}a_j \leq \tfrac{1}{2}|u_1| + \tfrac{1}{2}|u_2| = |u| = \lambda_u^{-1}a_j \qquad \text{on } |u|^{-1}\left(\lambda_u^{-1}a_j\right).$$

Now we conclude as in the first case that $\lambda_{u_i} = \lambda_u$ and thus $u_i = u$, so u is extreme.

Conversely suppose $u \in B_{\Psi,\mu}^{\circ\, e}$. Then $\|u\|_* = 1$ of course and if, for some j, (ii) fails then for some $\varepsilon > 0$ we have

$$\mu\left(|u|^{-1}\left[\lambda_u^{-1}a_j + \varepsilon, \lambda_u^{-1}b_j - \varepsilon\right]\right) > 0.$$

Exactly as in Theorem 3 we now partition

$$F = |u|^{-1}\left[\lambda_u^{-1}a_j + \varepsilon, \lambda_u^{-1}b_j - \varepsilon\right]$$

into disjoint subsets F_1 and F_2 of equal measure; because Φ is linear on $[a_j, b_j]$, for $|t| \leq \varepsilon$, $u_t = u + t(\chi_{F_1} - \chi_{F_2}) \operatorname{sgn} u$ will have

$$\int \Phi(\lambda_u |u_t|)\, d\mu = \int \Phi(\lambda_u |u|)\, d\mu$$

so that

$$(1/\lambda_u)\left(1 + \int \Phi(\lambda_u |u_t|)\, d\mu\right) = 1,$$

whence $\|u_t\|_* \leq 1$ by the final assertion of Lemma 2. So (ii) must hold.

Finally suppose (iii) fails, so $\mu(|u|^{-1}(\bigcup[\lambda_u^{-1}a_j, \lambda_u^{-1}b_j])) = 1$, $\int \Psi(\varphi(\lambda_u|u|))\, d\mu = 1$ and $\mu(|u|^{-1}(\lambda_u^{-1}a_j)) = 0$ for all j. Because of the first and last of these, and (ii), we have the range of $|u|$ in the sequence $\{\lambda_u^{-1}b_j\}$, and for $v = \varphi(\lambda_u|u|) \operatorname{sgn} u = \varphi(b_j) \operatorname{sgn} u$ on $|u|^{-1}[\lambda_u^{-1}a_j, \lambda_u^{-1}b_j]$, the second condition says $\int \Psi(|v|)\, d\mu = 1$. Of course $\int uv\, d\mu = 1$ since (19) holds.

Now by hypothesis $b_j = \lambda_u|u| \geq \eta a_j$ on $|v|^{-1}(c_j)$ for all j, with $\eta > 1$, so that if $\eta > \eta_1 > 1$ we can assert that for $\lambda = \eta_1^{-1}\lambda_u$ we have $b_j > \lambda|u| \geq \eta_1^{-1}\eta a_j > a_j$ on $|v|^{-1}(c_j)$. Thus $\int \Psi(\varphi(\lambda|u|))\, d\mu = 1$ still, as $\sum_j \Psi(c_j)\mu(|v|^{-1}(c_j))$. Indeed for some set $F \subset |v|^{-1}(c_j)$ of positive measure and $\varepsilon > 0$ small, clearly the same is true with u replaced by $u_\pm = u \pm \varepsilon(\chi_{F_1} - \chi_{F_2}) \operatorname{sgn} u$, where again $F_1 \cup F_2 = F$ and $\mu F_1 = \mu F_2$, and

$$\int u_\pm v\, d\mu = 1 = \int uv\, d\mu$$

since $|v| = c_j$ on F, and $\operatorname{sgn} u = \operatorname{sgn} v$. But now

$$\int \lambda u_\pm v\, d\mu = \int \Phi(\lambda|u_\pm|)\, d\mu + \int \Psi(|v|)\, d\mu = 1 + \int \Phi(\lambda|u_\pm|)\, d\mu,$$

so $1 = \int u_\pm v\, d\mu = (1/\lambda)(1 + \int \Phi(\lambda|u_\pm|)\, d\mu) \geq \|u_\pm\|_*$. Since $u = \frac{1}{2}(u_+ + u_-)$ we have our contradiction, showing (iii) cannot fail and completing our proof of Theorem 3°.

We can now simply observe that our adaptation of Lindenstrauss' proof to show $\rho(B^e_{\Phi,\mu}) = \rho(B_{\Phi,\mu})$ applies with little change to yield the assertion of Theorem 1 in the present case (where we assume $b_j/a_j \geq \theta > 1$). Indeed if u_0 is extreme in the subset of $B^\circ_{\Psi,\mu}$ mapping onto a given x_0 in $\mathbf{R}^n$, and not in $B^{\circ e}_{\Psi,\mu}$ then $\|u_0\|_* = 1$ (or choosing $u_1 \neq 0$ in kern ρ we have $u_0 \pm \varepsilon u_1$ in $B^\circ_{\Psi,\mu}$). Since u_0 is not extreme, by Theorem 3° we know that (iii) fails or that there is a j and $\varepsilon > 0$ for which $|u_0|^{-1}[\lambda_{u_0}^{-1}a_j + \varepsilon, \lambda_{u_0}^{-1}b_j - \varepsilon]$ has positive measure. In the second case we now choose a non-zero u supported by that set with $\operatorname{sgn} u = \operatorname{sgn} u_0$, $\rho(u) = 0$ and $(1/\lambda_{u_0})(1 + \int \Phi(\lambda_{u_0}|u_0 \pm u|)\, d\mu) = 1$, all of which are possible (with $|u| = \varepsilon$ on our set) by Liapounoff's theorem as before.

Just as in the preceding paragraph the last condition says $\|u_0 \pm u\|_* \leq 1$, contradicting the extremity of u_0 in $\rho^{-1}(x_0) \cap B^{\circ}_{\Psi,\mu}$.

Finally in case (iii) fails we have $|u_0| = \lambda_{u_0}^{-1} b_j$ on $|u_0|^{-1}[\lambda_{u_0}^{-1} a_j, \lambda_{u_0}^{-1} b_j]$ for all j and we can proceed as in the final paragraph of the proof of Theorem 3°, using $\lambda = \eta_1^{-1}\lambda_{u_0}$, to replace u_0 by $u = u_0 \pm \varepsilon u_1 \operatorname{sgn} u_0$, where $|u_1| \leq 1$ and ε is small enough to guarantee $\int \Psi(\varphi(\lambda|u|))\, d\mu = \sum \Psi(c_j)\mu(|v|^{-1}(c_j)) = 1$ as before, while $\int u_1 v\, d\mu = 0$ and $\int u_1 y\, d\mu = \rho(u_1) = 0$, $u_1 \neq 0$, is guaranteed by Liapounoff's theorem. Thus (iii) fails only if u_0 is not extreme, and our proof of Theorem 1 is complete.

In order to prove Theorem 2 for $B^{\circ}_{\Psi,\mu}$ we first note that by Lemma 2° for $x_0 = \rho(\chi_E u) \in \partial K_y(E)$ and θ supporting there $(\lambda_{\theta,u}|u(t)|, |\theta \cdot y(t)|)$ lies on the extended graph of φ a.e., where $\lambda_{\theta,\mu} = 1 + \int \Phi(\lambda_{\theta,\mu}|u|)\, d\mu \geq 1$, so one has $|u| \leq \|y\| = M$ with a bound independent of x_0. Again one has $K_y(E) = K_y^1(E)$, so the former is closed as before.

Once more we obtain our minimal E with $x_0 \in \partial\rho(\chi_E B^{\circ}_{\Psi,\mu})$. Now exactly as in the proof of the first assertion of Theorem 2 we can take $(y_1\chi_E, \ldots, y_n\chi_E)$ independent, and thus for some $\varepsilon > 0$

$$\mu\{t \in E: |\theta \cdot y(t)| > \varepsilon\} > \varepsilon \quad \text{for all } \theta \in S^{n-1}. \tag{24}$$

As a consequence, in $L_\Psi(\mu_E)$ $\|\theta \cdot y\|$ has a positive lower bound independent of θ: for if we choose $k > 0$ so that $\Psi(k) > 1/\varepsilon$ then by (24)

$$\int \Psi\left(\frac{k}{\varepsilon}|\theta \cdot y|\right) d\mu \geq \Psi(k) \cdot \varepsilon > 1$$

so $(k/\varepsilon)\|\theta \cdot y\| \geq 1$, and $m = \varepsilon/k$ is our lower bound. Thus

$$\left|\frac{\theta \cdot y(t)}{\|\theta \cdot y\|}\right| \leq m^{-1}\|y\| \leq m^{-1}M.$$

Let M_1 be the larger of $m^{-1}M$ and $\psi(m^{-1}M)$.

If we now alter φ on $(2M_1, \infty)$ to obtain a φ_1 yielding a dual pair Φ_1, Ψ_1 of Young's functions both satisfying the Δ_2-condition, as earlier, then we know $x_0 \in \partial\rho(\chi_E B^{\circ}_{\Psi,\mu}) = \partial\rho(\chi_E B^{\circ}_{\Psi_1,\mu})$ and we can now appeal to Theorem 1 to obtain $u \in B^{\circ e}_{\Psi_1,\mu_E}$ with $\rho(\chi_E u) = x_0$, and $|u| \leq M$ as above. Moreover for some $\theta \in S^{n-1}$, $\chi_E\theta \cdot y$ maximizes over $B^{\circ}_{\Psi_1,\mu_E}$ at u, and with v the normalized element of $L_{\Psi_1}(\mu_E)$ $\theta \cdot y/\|\theta \cdot y\|$ (so $\int \Psi_1(|v|)\, d\mu_E = 1$) we know by Lemma 2° that $(\lambda_u^1|u|, |v|)$ lies a.e. on the extended graph of φ_1, (hence of φ since $\lambda^1|u| \leq \psi(m^{-1}M) \leq M_1$) where

$$\lambda_u^1 = \sup\left\{\lambda > 0: \int \Psi_1\left(\varphi_1(\lambda|u|)\right) d\mu \leq 1\right\}. \tag{24}$$

Further (still in terms of $L_{\Phi_1}(\mu_E)$)

$$1 = \|u\|_* = (1/\lambda_u^1)\left(1 + \int \Phi_1(\lambda_u^1|u|)\, d\mu_E\right),$$

and so $\int uv\, d\mu_E = 1$ (essentially since (19) obtains):

$$\int \lambda_u^1 uv\, d\mu_E = \int \Phi_1(\lambda_u^1|u|)\, d\mu_E + \int \Psi_1(|v|)\, d\mu_E$$

$$= \int \Phi_1(\lambda_u^1|u|)\, d\mu_E + 1 = \lambda_u^1.$$

But $1 = \int uv\, d\mu_E = \int \Psi_1(|v|)\, d\mu_E$ implies $\int \Psi(|v|)\, d\mu_E = 1$, whence $\|v\| = 1$ in $L_\Psi(\mu_E)$, and consequently that u has its functional norm 1 in $L_\Phi(\mu_E)$. Finally we claim λ_u^1 in (24) coincides with

$$\lambda_u = \sup\left\{\lambda > 0\colon \int \Psi(\varphi(\lambda|u|))\, d\mu_E \le 1\right\}.$$

Indeed

$$1 = \frac{1}{\lambda_u^1}\left(1 + \int \Phi_1(\lambda_u^1|u|)\, d\mu_E\right) = \frac{1}{\lambda_u^1}\left(1 + \int \Phi(\lambda_u^1|u|)\, d\mu_E\right)$$

since $\lambda_u^1|u| \le \psi(m^{-1}M) \le M_1$, and thus λ_u^1 provides

$$\min_{\lambda > 0} \frac{1}{\lambda}\left(1 + \int \Phi(\lambda|u|)\, d\mu_E\right) = \|u\|_* = 1.$$

But from our remark following Lemma 2 our minimum is provided only on an interval $[\lambda_u - \varepsilon, \lambda_u]$ (so $\lambda_u^1 \le \lambda_u$), and then for λ in that interval $\lambda|u|$ has its range entirely in our maximal intervals of constancy:

$$\mu_E\left(|u|^{-1}\left(\bigcup\left[\lambda^{-1}a_j, \lambda^{-1}b_j\right]\right)\right) = 1, \tag{25}$$

and $\int \Psi(\varphi(\lambda|u|))\, d\mu_E = 1$ (because the derivative in our remark is zero on $[\lambda_u - \varepsilon, \lambda_u)$). Thus from the converse portion of Theorem 3°(iii) for $\lambda = \lambda_u^1$ we know $|u| = (\lambda_u^1)^{-1}a_j$ on a set of positive measure for some j, so that (25) cannot hold for all λ in $(\lambda_u^1, \lambda_u^1 + \delta)$, $\delta > 0$. Consequently $\lambda_u^1 < \lambda_u$ is impossible, and $\lambda_u^1 = \lambda_u$ as asserted.

But now $u \in B^{\circ e}_{\Psi_1, \mu_e}$ has its consequences from Theorem 3° precisely the conditions sufficient to insure $u \in B^{\circ e}_{\Psi, \mu_e}$ by the same result. If φ has no initial interval of constancy we can simply extend u to be 0 on $X \setminus E$ obtaining the desired element of $B^{\circ e}_{\Psi, \mu}$ (by Theorem 3° again) mapping onto x_0. But if we have an initial interval of constancy $[0, b]$ we can choose an element u_1 of the b ball of $L^\infty(\mu_{X\setminus E})$ with $\rho(\chi_{X \setminus E} u_1) = 0$ and

$|u| \equiv b$ by Liapounoff's theorem, so $u + u_1$ is evidently the desired element of $B^{\circ e}_{\Psi,\mu}$. Our proof of Theorem 2 is finally complete.

Lastly we note that complex versions of these results only require appropriate attention to conjugation and absolute values at various points in the proofs.

REFERENCES

[1] Z. Birnbaum and W. Orlicz, *Über die Veralgemeinerung des Begriffes derzueinander konjugiertan Potentzen*, Studia Math., **3** (1931), 1–67.

[2] I. Glicksberg, *A controlled approach to Liapounoff's convexity theorem*, J. Math. Anal. and Appl., **94** (1983), 193–196.

[3] M. A. Kransnoselskii and Y. B. Rutickii, *Convex Functions and Orlicz Spaces*, P. Noordhoff Ltd, Groningen, 1961.

[4] A. Liapounoff, *Sur les fonctions-vecteurs complétement additives*, Bull. Acad. Sci. URSS, **4** (1940), 465–478.

[5] J. Lindenstrauss, *A short proof of Liapounoff's convexity theorem*, J. Math. Mech., **15** (1966), 971–2.

[6] W. Rudin, *Functional Analysis*, McGraw Hill, N.Y., 1973.

[7] A. Zygmund, *Trigonometric Series*, 2nd ed., Cambridge Univ. Press, 1958.

Received May 25, 1983. Work supported in part by the NSF.

UNIVERSITY OF WASHINGTON
SEATTLE, WA 98195