Convex Geometric Analysis
MSRI Publications
Volume **34**, 1998

Polytope Approximations of the Unit Ball of ℓ_p^n

W. TIMOTHY GOWERS

Abstract. A simple and explicit method is given for approximating the unit ball of ℓ_p^n by polytopes. The method leads to a natural generalization of ℓ_p-spaces with good duality and interpolation properties.

1. Introduction

The classical spaces ℓ_p and L_p are the best known and in many ways most fundamental examples of Banach spaces. In view of their interesting properties it is natural to ask whether the role of the function t^p in these spaces can be played by other more general functions. This question was answered by Orlicz, who defined a certain class of functions, now known as Orlicz functions, and associated with each one a sequence space and a function space, now called an Orlicz sequence space and Orlicz function space. The Orlicz spaces are generally regarded as the correct and most natural spaces to associate with given Orlicz functions.

One of the aims of this paper is to cast doubt on that view, at least in its isometric interpretation. We shall do this by discussing a different generalization which arises geometrically and has two desirable isometric properties lacked by Orlicz spaces. First, the dual of one of our spaces is isometric to another such space. Second, complex interpolation between two of our spaces yields a third in a natural way. Irritatingly, we have not managed to establish whether our new spaces are *isomorphic* to Orlicz spaces, in which case they are a useful renorming of them, or whether they are completely different. Our route to the new generalization starts with an unusual (perhaps even eccentric) problem which will be described below, and which relates more to the polytope approximations of the title.

The results of this paper originated with the following line of thought. A common way of constructing a finite-dimensional normed space which lacks symmetry properties, invented by Gluskin [G], is to take a small number of antipodal pairs of points at random from the unit sphere of ℓ_2^n and to take their convex hull as the unit ball of the space. Alternatively, and dually, one can take the convex polytope defined by the hyperplanes tangent to the sphere at the given points. However, such spaces have properties which a general space cannot be expected to have. For example, geometrically their unit balls are far from being typical centrally symmetric polytopes, having in the first case very few vertices and in the second case very few faces. A related fact is that the convex hull of a small number of points gives a space with very small cotype constants. Is there some way of constructing a convex body which is in any useful sense completely generic?

We shall return to this question at the end of the paper, but it is not our main concern. Instead, we shall give an (incomplete) investigation of what happens when one mixes the process of taking the convex hull of a few points with the dual process. One is led naturally away from generic spaces and towards very special ones. At the end of the paper we make several suggestions for how the investigation might be continued.

2. A Polytope Approximation of the Unit Ball of ℓ_2^n

Recall that a basis $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ of an n-dimensional normed space is said to be 1-*symmetric* if, for every choice of scalars $a_1, \dots, a_n$, every choice of signs $\varepsilon_1, \dots, \varepsilon_n$ with $\varepsilon_i = \pm 1$, and every permutation π of $\{1, 2, \dots, n\}$, we have the equality

$$\left\| \sum_{i=1}^n \varepsilon_i a_i \boldsymbol{e}_{\pi(i)} \right\| = \left\| \sum_{i=1}^n a_i \boldsymbol{e}_i \right\|.$$

For $n \geqslant 1$ let $\mathcal{F}_n$ stand for the set of normed spaces of the form $(\mathbb{R}^n, \|\cdot\|)$ for which the standard basis $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ is normalized and 1-symmetric. We also define two operations S and T which map $\bigcup_{n=1}^\infty \mathcal{F}_n$ to $\bigcup_{n=1}^\infty \mathcal{F}_{n+1}$ as follows. Given $X \in \mathcal{F}_n$, we set $T(X)$ to be the unique normed space in $\mathcal{F}_{n+1}$ such that for any $\boldsymbol{a} = \sum_1^{n+1} a_i \boldsymbol{e}_i \in \mathbb{R}^{n+1}$ with $a_1 \geqslant \dots \geqslant a_{n+1} \geqslant 0$, $\|\boldsymbol{a}\|_{T(X)} = \| \sum_1^n a_i \boldsymbol{e}_i \|_X$. The norm on $T(X)$ is as small as it can be given that it is 1-symmetric and that its restriction to any n coordinates yields the norm on X. The operation S is defined in the opposite way: the norm on $S(X)$ is as big as it can be under the same conditions. Thus $S(X) = (T(X^*))^*$. Alternatively, given the unit ball of X, embed it into $\mathbb{R}^{n+1}$ in the $n+1$ natural ways (up to symmetry) given by ignoring each of the coordinates in turn. The convex hull of these $n+1$ copies is the unit ball of $S(X)$. Note that the unit ball of $T(X)$ has a geometric description as well. The Cartesian product of one of the $n+1$ images of the unit ball of X with the direction not used is a cylinder. The unit ball of $T(X)$ is the intersection of these $n+1$ cylinders.

Now, if we start with the single space in $\mathcal{F}_1$, the normalized 1-dimensional space, which we shall call $\mathbb{R}$, and apply S $n-1$ times, we clearly end up with ℓ_1^n. Similarly, if we apply T $n-1$ times we obtain ℓ_∞^n. What happens if we alternate

S with T? The answer is the main result of this section and the motivation for the rest of the paper.

THEOREM 1. *Let $n \geqslant 2$ and let X be a space in $\mathcal{F}_n$ obtained from the space $\mathbb{R}$ by applying the operations S and T alternately. Then $d(X, \ell_2^n) = \sqrt{2}$.*

The proof of this theorem is based on two lemmas: the first is trivial and the second also easy.

LEMMA 2. *If $X, Y \in \mathcal{F}_n$ and $\|\boldsymbol{a}\|_X \leqslant \|\boldsymbol{a}\|_Y$ for every $\boldsymbol{a} \in \mathbb{R}^n$, then $\|\boldsymbol{a}\|_{SX} \leqslant \|\boldsymbol{a}\|_{SY}$ and $\|\boldsymbol{a}\|_{TX} \leqslant \|\boldsymbol{a}\|_{TY}$ for every $\boldsymbol{a} \in \mathbb{R}^{n+1}$.*

LEMMA 3. *Suppose $X = \ell_2^n$. Then $\|\boldsymbol{a}\|_{\ell_2^{n+2}} \leqslant \|\boldsymbol{a}\|_{TSX}$ for every $\boldsymbol{a} \in \mathbb{R}^{n+2}$.*

Before proving Lemma 3, let us see why the two lemmas are sufficient to prove the main result. In the argument that follows and for the rest of the paper it will be convenient, when $X, Y \in \mathcal{F}_n$, to use the abbreviation $c_1 X \leqslant c_2 Y$ for the statement that $c_1\|\boldsymbol{a}\|_X \leqslant c_2\|\boldsymbol{a}\|_Y$ for any vector $\boldsymbol{a} \in \mathbb{R}^n$.

PROOF OF THEOREM 1. First consider the space $(TS)^k(\mathbb{R})$ for some $k \in \mathbb{N}$. It is clear from the two lemmas that $(TS)^k(\mathbb{R}) \geqslant \ell_2^{2k+1}$. Moreover, since $T(\mathbb{R}) \geqslant 2^{-1/2}\ell_2^2$, we also have that $(TS)^k T(\mathbb{R}) \geqslant 2^{-1/2}\ell_2^{2k+2}$, and hence that $S(TS)^{k-1}T(\mathbb{R}) = (ST)^k(\mathbb{R}) \geqslant 2^{-1/2}\ell_2^{2k+1}$. However, $(TS)^k(\mathbb{R})$ and $(ST)^k(\mathbb{R})$ are dual to each other, so we have the relations

$$2^{-1/2}\ell_2^{2k+1} \leqslant (ST)^k(\mathbb{R}) \leqslant \ell_2^{2k+1} \leqslant (TS)^k(\mathbb{R}) \leqslant 2^{1/2}\ell_2^{2k+1}.$$

It follows immediately also that

$$2^{-1/2}\ell_2^{2k} \leqslant T(ST)^{k-1}(\mathbb{R}) \leqslant \ell_2^{2k} \leqslant S(TS)^{k-1}(\mathbb{R}) \leqslant 2^{1/2}\ell_2^{2k}.$$

This establishes that $d(X, \ell_2^n) \leqslant \sqrt{2}$, conditional on the truth of Lemma 3. It is well known that the distance from a space in $\mathcal{F}_n$ to ℓ_2^n is attained by the identity map. Considering the norms in X and ℓ_2^n of $\boldsymbol{e}_1$ and $\boldsymbol{e}_1 + \boldsymbol{e}_2$, we obtain the reverse inequality. $\square$

PROOF OF LEMMA 3. Throughout this proof, all sequences will be assumed to be positive and decreasing. Writing $X = \ell_2^n$ we have, for any $\boldsymbol{a} \in \mathbb{R}^{n+1}$ and $\boldsymbol{b} \in \mathbb{R}^{n+2}$,

$$\|\boldsymbol{a}\|_{SX} = \max\left\{ \sum_1^{n+1} f_i a_i : \sum_1^n f_i^2 \leqslant 1 \right\}$$

and

$$\|\boldsymbol{b}\|_{TSX} = \max\left\{ \sum_1^{n+1} f_i b_i : \sum_1^n f_i^2 \leqslant 1 \right\}.$$

Hence it is enough to show that, for any $a_1 \geqslant \cdots \geqslant a_{n+2} \geqslant 0$ with $\sum_1^{n+2} a_i^2 = 1$, we can find $f_1 \geqslant \cdots \geqslant f_{n+1} \geqslant 0$ such that $\sum_1^n f_i^2 \leqslant 1$ and $\sum_1^{n+1} f_i a_i \geqslant 1$. We do this by setting $f_i = \lambda a_i$ for $1 \leqslant i \leqslant n$, where $\lambda = (1 - a_{n+1}^2 - a_{n+2}^2)^{-1/2}$, and setting $f_{n+1} = f_n$. Then certainly $\sum_1^n f_i^2 = 1$, and moreover

$$\sum_1^{n+1} f_i a_i = \lambda \sum_1^n a_i^2 + \lambda a_n a_{n+1} = \frac{1 - a_{n+1}^2 - a_{n+2}^2 + a_n a_{n+1}}{(1 - a_{n+1}^2 - a_{n+2}^2)^{1/2}}$$
$$\geqslant \frac{1 - a_{n+2}^2}{(1 - 2a_{n+2}^2)^{1/2}} \geqslant 1. \qquad \square$$

The proof of Theorem 1 can be generalized very easily to approximate ℓ_p^n in a similar way when p or its conjugate is an integer. However, with a little more work, one can approximate ℓ_p^n for an arbitrary p. We shall do this in Section 4. First, we show how to calculate the norm of a vector in a space of the form $U_1(U_2(\dots U_{n-1}(\mathbb{R})\dots))$ where each U_i is either S or T. We shall call such a space an ST-space.

3. An Algorithm for Calculating the Norm

Let $X \in \mathcal{F}_n$ and let $\boldsymbol{a} = \sum_1^{n+1} a_i \boldsymbol{e}_i$ be a vector in $\mathbb{R}^{n+1}$ such that $a_1 \geqslant \dots \geqslant a_{n+1} \geqslant 0$. Then by definition $\|\boldsymbol{a}\|_{TX} = \|\sum_1^n a_i \boldsymbol{e}_i\|_X$. We shall now show how to calculate $\|\boldsymbol{a}\|_{SX}$. Given two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ in $\mathbb{R}^n$, write $\boldsymbol{a}^* = (a_1^*, \dots, a_n^*)$ and $\boldsymbol{b}^* = (b_1^*, \dots, b_n^*)$ for their (non-negative) decreasing rearrangements. Write $\boldsymbol{a} \prec \boldsymbol{b}$ if $\sum_{i=1}^k a_i^* \leqslant \sum_{i=1}^k b_i^*$ for every $k \leqslant n$. It is well known that $\boldsymbol{a} \prec \boldsymbol{b}$ if and only if $\boldsymbol{a}$ lies in the convex hull of all vectors that can be obtained from $\boldsymbol{b}$ by permuting the coordinates and changing some of their signs. Therefore, if $\|\cdot\|$ is a 1-symmetric norm and $\boldsymbol{a} \prec \boldsymbol{b}$, we know that $\|\boldsymbol{a}\| \leqslant \|\boldsymbol{b}\|$.

Given $X \in \mathcal{F}_n$ and $\boldsymbol{a} \in \mathbb{R}^{n+1}$, we define a function $\|\cdot\|'$ on $\mathbb{R}^{n+1}$ by

$$\|\boldsymbol{a}\|' = \min\big\{\|\boldsymbol{b}\|_X : \boldsymbol{b} \in X,\ \boldsymbol{a} \prec \boldsymbol{b}\big\},$$

where $\boldsymbol{a} \prec \boldsymbol{b}$ means of course that $\boldsymbol{a} \prec \boldsymbol{b}'$, where $\boldsymbol{b}'$ is the image of $\boldsymbol{b}$ in $\mathbb{R}^{n+1}$ under the obvious inclusion. (Note that the existence of the above minimum follows easily from compactness.)

LEMMA 4. *Let $X \in \mathcal{F}_n$ and $\boldsymbol{a} \in \mathbb{R}^{n+1}$. Then $\|\boldsymbol{a}\|_{SX} = \|\boldsymbol{a}\|'$.*

PROOF. Given two vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ in $\mathbb{R}^{n+1}$, let $\boldsymbol{b}_i \in \mathbb{R}^n$ be such that $\boldsymbol{a}_i \prec \boldsymbol{b}_i$ and $\|\boldsymbol{b}_i\|_X$ is minimal, for $i = 1, 2$. Then we certainly have $\boldsymbol{a}_1 + \boldsymbol{a}_2 \prec \boldsymbol{b}_1^* + \boldsymbol{b}_2^*$, and thus, since X is a 1-symmetric space,

$$\|\boldsymbol{a}_1 + \boldsymbol{a}_2\|' \leqslant \|\boldsymbol{b}_1^* + \boldsymbol{b}_2^*\|_X \leqslant \|\boldsymbol{b}_1^*\|_X + \|\boldsymbol{b}_2^*\|_X = \|\boldsymbol{b}_1\|_X + \|\boldsymbol{b}_2\|_X = \|\boldsymbol{a}_1\|' + \|\boldsymbol{a}_2\|'.$$

It follows that $\|\cdot\|'$ is a norm. By the discussion above concerning the relation $\prec$, the unit ball of the space $(\mathbb{R}^{n+1}, \|\cdot\|')$ is contained in that of SX. Since it also contains the $n+1$ natural images of $B(X)$ in $\mathbb{R}^{n+1}$, we have that $\|\cdot\|' = \|\cdot\|_{SX}$, as stated. $\square$

In the next lemma we identify a vector at which the minimum of the set

$$\big\{\|\boldsymbol{b}\|_X : \boldsymbol{b} \in X,\, \boldsymbol{a} \prec \boldsymbol{b}\big\}$$

is attained. We write $x \vee y$ for $\max\{x, y\}$.

Lemma 5. *Let $X \in \mathcal{F}_n$, let $\boldsymbol{a} = \sum_1^{n+1} a_i \boldsymbol{e}_i$ be a vector in $\mathbb{R}^{n+1}$ for which $a_1 \geqslant \cdots \geqslant a_{n+1} > 0$ and let $\gamma \geqslant 0$ be the unique number for which*

$$\sum_1^n (a_i \vee \gamma) = \sum_1^{n+1} a_i.$$

Then $\|\boldsymbol{a}\|_{SX} = \big\|\sum_1^n (a_i \vee \gamma)\boldsymbol{e}_i\big\|_X$.

Proof. We insist that $a_{n+1} > 0$ for convenience: it is obvious how to calculate $\|\boldsymbol{a}\|_{SX}$ if $a_{n+1} = 0$. This gives us the uniqueness of γ. Set $\boldsymbol{a}' = \sum_1^n (a_i \vee \gamma)\boldsymbol{e}_i$. We need to show that if $\boldsymbol{b} \in \mathbb{R}^n$ and $\boldsymbol{a} \prec \boldsymbol{b}$ then $\boldsymbol{a}' \prec \boldsymbol{b}$. It is clear that $\boldsymbol{a} \prec \boldsymbol{a}'$, and by our earlier remarks and the symmetry of X, we will also have $\|\boldsymbol{a}'\|_X \leqslant \|\boldsymbol{b}\|_X$. We may clearly suppose that $b_i = b_i^*$ for every i and that $\sum_1^n b_i = \sum_1^{n+1} a_i$. Since $\boldsymbol{a} \prec \boldsymbol{b}$ we then have, for every $1 \leqslant k \leqslant n$, that $\sum_k^n b_i \leqslant \sum_k^{n+1} a_i$. Let k be maximal such that $a_k > \gamma$. Clearly $k < n$, and $\gamma = (n-k)^{-1} \sum_{k+1}^{n+1} a_i$. For $l \leqslant k$ it is obvious that $\sum_1^l (a_i \vee \gamma) \leqslant \sum_1^l b_i$. When $l > k$ we have

$$\begin{aligned}\sum_1^l (a_i \vee \gamma) = \sum_1^{n+1} a_i - (n-l)\gamma &= \sum_1^{n+1} a_i - \frac{n-l}{n-k}\sum_{k+1}^{n+1} a_i \\ &\leqslant \sum_1^{n+1} a_i - \frac{n-l}{n-k}\sum_{k+1}^{n} b_i \leqslant \sum_1^{n+1} a_i - \sum_{l+1}^{n} b_i = \sum_1^{l} b_i.\end{aligned}$$

This proves the lemma. □

To conclude this section we shall show how to calculate the norm of the vector $(3, 3, 3, 2, 2, 2, 1, 1, 1)$ in the spaces $(ST)^4(\mathbb{R})$ and $(TS)^4(\mathbb{R})$ by repeated application of Lemma 5 and the definition of the operation T. At each stage, we replace the vector we have by one of length one less, while preserving its norm. At a "T" stage, we simply remove the last coordinate. At an "S" stage, we apply Lemma 5. The process is summarized in the table at the top of the next page.

Notice that $\|(3, 3, 3, 2, 2, 2, 1, 1, 1)\|_2 = \sqrt{42}$, and, as must be the case by the proof of Theorem 1, $\sqrt{21} \leqslant 5\frac{1}{4} \leqslant \sqrt{42} \leqslant 8 \leqslant \sqrt{84}$.

4. Approximating ℓ_p^n by an ST-Space

In order to show that ℓ_p^n can be approximated by an ST-space, it turns out to be convenient and natural to generalize the notion of ST-space to function spaces (although these do not appear in the final statement). We shall prove an inequality which is a little more sophisticated than Lemma 3. However, our main difficulty is notational rather than conceptual. In the next section, we shall

$$
\begin{array}{cccccccccc}
 & 3 & 3 & 3 & 2 & 2 & 2 & 1 & 1 & 1 \\
\xrightarrow{S} & 3 & 3 & 3 & 2 & 2 & 2 & 1\frac{1}{2} & 1\frac{1}{2} & \\
\xrightarrow{T} & 3 & 3 & 3 & 2 & 2 & 2 & 1\frac{1}{2} & & \\
\xrightarrow{S} & 3 & 3 & 3 & 2\frac{1}{2} & 2\frac{1}{2} & 2\frac{1}{2} & & & \\
\xrightarrow{T} & 3 & 3 & 3 & 2\frac{1}{2} & 2\frac{1}{2} & & & & \\
\xrightarrow{S} & 3\frac{1}{2} & 3\frac{1}{2} & 3\frac{1}{2} & 3\frac{1}{2} & & & & & \\
\xrightarrow{T} & 3\frac{1}{2} & 3\frac{1}{2} & 3\frac{1}{2} & & & & & & \\
\xrightarrow{S} & 5\frac{1}{4} & 5\frac{1}{4} & & & & & & & \\
\xrightarrow{T} & 5\frac{1}{4} & & & & & & & &
\end{array}
\qquad
\begin{array}{cccccccccc}
 & 3 & 3 & 3 & 2 & 2 & 2 & 1 & 1 & 1 \\
\xrightarrow{T} & 3 & 3 & 3 & 2 & 2 & 2 & 1 & 1 & \\
\xrightarrow{S} & 3 & 3 & 3 & 2 & 2 & 2 & 2 & & \\
\xrightarrow{T} & 3 & 3 & 3 & 2 & 2 & 2 & & & \\
\xrightarrow{S} & 3 & 3 & 3 & 3 & 3 & & & & \\
\xrightarrow{T} & 3 & 3 & 3 & 3 & & & & & \\
\xrightarrow{S} & 4 & 4 & 4 & & & & & & \\
\xrightarrow{T} & 4 & 4 & & & & & & & \\
\xrightarrow{S} & 8 & & & & & & & &
\end{array}
$$

explain a different and in some ways more satisfactory approach, which shows again that ℓ_p^n can be approximated by an ST-space, while avoiding the use of a technical lemma from this section (Lemma 7 below).

Given a real number $t > 0$, let $F[0,t]$ denote the vector space of step functions from the closed interval $[0,t]$ to $\mathbb{R}$, and let $\mathcal{G}_t$ denote the set of normed spaces $(F[0,t], \|\cdot\|)$ such that $\|f\| = \|f^*\|$ for any $f \in F[0,t]$, where f^* is the decreasing rearrangement of f. Note that we do not ask for these normed spaces to be complete. If $s > t$ and $f \in F[0,s]$ we shall write f_t for the restiction of the function f to the interval $[0,t]$. For $n \in \mathbb{N}$ let I_n denote the linear map from $\mathbb{R}^n$ to $F[0,n]$ determined by $e_j \mapsto \chi_{[j-1,j)}$. We shall prove a result about norms on $F[0,t]$ but our interest will eventually be in subspaces of the form $I_n(\mathbb{R}^n)$.

We now define operations which are similar to S and T, but which map $\mathcal{G}_{t_1}$ to $\mathcal{G}_{t_2}$, where $t_1 < t_2$. Given $\alpha > 1$, define an operation $T_\alpha : \bigcup_{t>0} \mathcal{G}_t \longrightarrow \bigcup_{t>0} \mathcal{G}_{\alpha t}$ as follows. If $X \in \mathcal{G}_t$ and $f \in F[0,\alpha t]$ then $\|f\|_{T_\alpha(X)} = \|(f^*)_t\|_X$. Thus, the norm on the space $T_\alpha(X)$ is as small as it can be given that $T_\alpha(X)$ is in the set $\mathcal{G}_{\alpha t}$ and that the norm on $T_\alpha(X)$ coincides with that on X for functions supported on the interval $[0,t]$. As in the discrete case, S_α is defined in the opposite way: the norm on $S_\alpha(X)$ is as large as it can be under the same conditions. The following lemma we state without proof, since the analogy with Lemma 5 is very close.

Lemma 6. *Let $t > 0$, $\alpha > 1$ and $X \in \mathcal{G}_t$. Let $f \in F[0,\alpha t]$ be a step function satisfying $f = f^*$ and $f(s) > 0$ for some $s > t$. Then if $\gamma > 0$ is the unique number for which*

$$\int_0^t (f(t) \vee \gamma)\, dt = \int_0^{\alpha t} f(t)\, dt\,,$$

we have $\|f\|_{S_\alpha(X)} = \|(f \vee \gamma)_t\|_X$.

We now come to the main lemma of this section. As in the first section, if X and Y are spaces in $\mathcal{G}_t$ and c_1 and c_2 are positive constants, then we shall write $c_1 X \leqslant c_2 Y$ if $c_1 \|f\|_X \leqslant c_2 \|f\|_Y$ for every $f \in F[0,t]$. For $1 \leqslant p \leqslant \infty$ and $t > 0$

let L_p^t denote the space $(F[0,t], \|\cdot\|_p)$, where $\|\cdot\|_p$ is the usual norm on $L_p[0,t]$ restricted to $F[0,t]$.

LEMMA 7. *Let $1 < p < \infty$, $t > 0$, $\alpha > 1$ and $\beta = \alpha^{p-1}$. Then*

$$S_\alpha(T_\beta(L_p^t)) \leqslant L_p^{\alpha\beta t} \leqslant T_\beta(S_\alpha(L_p^t)).$$

PROOF. We shall prove the right-hand inequality only. The left-hand one can be proved by a similar argument, or else by using duality. Suppose then that the right-hand inequality does not hold. In that case we can find $N \in \mathbb{N}$ and sequences $0 = x_0 < x_1 < \cdots < x_N = \alpha\beta t$ and $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_N \geqslant 0$ such that, setting $A_i = [x_{i-1}, x_i)$ and $f = \sum_1^N \lambda_i \chi_{A_i}$, we have

$$\int_0^{\alpha\beta t} f(x)^p\,dx > \int_0^t (f(x) \vee \gamma)^p\,dx$$

where

$$\int_0^t (f(x) \vee \gamma)\,dt = \int_0^{\alpha t} f(x)\,dx.$$

Let i_0 be the minimal index i for which $\lambda_i \leqslant \gamma$ and set $s = x_{i_0-1}$. Thus $f(x) \leqslant \gamma$ if and only if $x \geqslant s$. Without loss of generality, there exists i_1 such that $x_{i_1} = \alpha t$. We have

$$\gamma(t-s) = \int_s^{\alpha t} f(x)\,dx,$$

and, by assumption,

$$\int_0^{\alpha\beta t} f(x)^p\,dx - \int_0^t (f(x) \vee \gamma)^p\,dx$$
$$= \int_s^{\alpha\beta t} f(x)^p\,dx - (t-s)^{-p}\left(\int_s^{\alpha t} f(x)\,dx\right)^p (t-s) > 0.$$

Write A for $(t-s)^{p-1}\int_s^{\alpha\beta t} f(x)^p\,dx$ and B for $\left(\int_s^{\alpha t} f(x)\,dx\right)^p$. Our hypothesis now reads that $A > B$. If we fix the sequence $x_0, \ldots, x_N$, we may assume, by a simple compactness argument, that $\lambda_1 \geqslant \cdots \geqslant \lambda_N \geqslant 0$ have been chosen so as to maximize the ratio A/B. Pick $i_0 \leqslant i < i_1$. We find

$$\frac{\partial}{\partial\lambda_i}\frac{A}{B} = \frac{1}{B^2}\left(B(t-s)^{p-1}(x_i - x_{i-1})p\lambda_i^{p-1} - Ap\left(\int_s^{\alpha t} f(x)\,dx\right)^{p-1}(x_i - x_{i-1})\right)$$

But since $\lambda_i \leqslant \gamma$, we have also $\left(\int_s^{\alpha t} f(x)\,dx\right)^{p-1} \geqslant (t-s)^{p-1}\lambda_i^{p-1}$. Since $A > B$ we obtain that $\frac{\partial}{\partial\lambda_i}\frac{A}{B} < 0$. A simple calculation shows that if we decrease λ_i very slightly, then the value of i_0 does not change. If the change to λ_i is small enough, and the resulting sequence is still a decreasing one, then it follows that A/B was not maximized by the original sequence. Hence, we obtain that $\lambda_{i_0} = \lambda_{i_0+1} = \cdots = \lambda_{i_1}$. It is obvious also that $\lambda_{i_1} = \cdots = \lambda_N$. Without loss of generality they all take the value 1.

It remains to show that

$$(\alpha\beta t - s)(t-s)^{p-1} \leqslant (\alpha t - s)^p$$

By applying the weighted arithmetic-geometric-mean inequality twice and using the fact that $\beta = \alpha^{p-1}$, we have, writing q for the conjugate index of p,

$$\begin{aligned}(\alpha\beta t - s)^{1/p}(t-s)^{1/q} &= \alpha t\left(1 - \frac{s}{\alpha\beta t}\right)^{1/p}\left(1-\frac{s}{t}\right)^{1/q} \\ &\leqslant \alpha t\left(1 - \frac{s}{p\alpha\beta t} - \frac{s}{qt}\right) \leqslant \alpha t\left(1 - \frac{s}{\alpha t}\right) = \alpha t - s.\end{aligned}$$

The result follows on raising both sides to the power p. □

The next lemma is similar to Lemma 7, but rather easier.

LEMMA 8. *Let $t > 0$, $X \in \mathcal{G}_t$ and $\alpha, \beta > 1$. Then $S_\alpha T_\beta X \leqslant T_\beta S_\alpha X$.*

PROOF. Let $f \in F[0, \alpha\beta t]$. Without loss of generality $f = f^*$. Then $\|f\|_{S_\alpha T_\beta X} = \|(f \vee \gamma)_t\|_X$, where γ satisfies

$$\int_0^{\beta t} (f(x) \vee \gamma)\, dx = \int_0^{\alpha\beta t} f(x)\, dx$$

and $\|f\|_{T_\beta S_\alpha X} = \|(f \vee \gamma')_t\|_X$, where γ' satisfies

$$\int_0^{t} (f(x) \vee \gamma')\, dx = \int_0^{\alpha t} f(x)\, dx.$$

It is therefore enough to show that $\int_0^t (f(x) \vee \gamma)\, dx \leqslant \int_0^{\alpha t} f(x)\, dx$, and thus that $\gamma' \geqslant \gamma$. Let $s = \inf\{x : f(x) \leqslant \gamma\}$. If $s \geqslant t$, the result is trivial. Otherwise, we wish to show that $(t-s)\gamma \leqslant \int_s^{\alpha t} f(x)\, dx$, given that $\gamma = (\beta t - s)^{-1} \int_s^{\alpha\beta t} f(x)\, dx$. But since f is a positive decreasing function and α, $\beta > 1$, we have

$$\frac{\int_s^{\alpha\beta t} f(x)\, dx}{\int_s^{\alpha t} f(x)\, dx} \leqslant \frac{\alpha\beta t - s}{\alpha t - s} \leqslant \frac{\beta t - s}{t - s},$$

which proves the lemma. □

In fact, it is not hard to see that equality in the last line can occur only for functions supported on the interval $[0, t]$, or functions whose modulus is constant (except on a set of measure zero) on the interval $[0, \alpha\beta t]$.

Lemmas 7 and 8 contain the essence of the proof that we can approximate ℓ_p^n by an ST-space. The remaining arguments are easy, but this may not be immediately apparent. The reader is strongly advised to consider, for any space $X \in \mathcal{G}_t$ under discussion, a logarithmic graph of the function $\lambda(t) = \|\chi_{[0,t]}\|_X$, i.e., a graph of $\log \lambda$ against $\log t$. Note that the slope of this graph, whenever X is an ST-space, is either zero or one, and when $X = L_p^t$ it is $1/p$. The terminology that follows is needed in order to formalize arguments which from such a graph are simple.

Define an ST-*sequence* to be a sequence of the form $U_1, \dots, U_k$, where each U_i is S_α or T_α for some $\alpha = \alpha_i$. Given $0 < u < t$ we shall say that a function $\lambda : [u,t] \longrightarrow \mathbb{R}$ is an ST-*function* if it is piecewise linear and its right derivative at any x is either $\lambda(x)/x$ or zero. Given a space $X \in \mathcal{G}_u$, we shall say that a space Y is an ST-extension of X if $Y = U_k(\dots(U_1(X))\dots)$ for some ST-sequence $U_1, \dots, U_k$. Given a space $X \in \mathcal{G}_u$, there is an obvious one-to-one correspondence between ST-sequences, ST-extensions of X and ST-functions taking the value $\|\chi_{[0,u]}\|_X$ at u. (Given an ST-extension Y of X, the associated ST-function is the function $\lambda : x \mapsto \|\chi_{[0,x]}\|_Y$.) When we refer to any correspondence between ST-functions, ST-extensions and ST-sequences, we shall always mean this one. Finally, we shall call a function $\lambda : (0,t] \to \mathbb{R}$ an ST-function if its restriction to $[u,t]$ is an ST-function for every $u > 0$. In other words, the conditions are as above except that there may be infinitely many changes of slope near zero. We associate with such a function a space as follows. Suppose the changes of slope of λ occur at points $\dots, x_{-2}, x_{-1}, x_0, x_1, \dots, x_n$ with $x_n < t$ and $x_{-i} \to 0$. Given a step function f defined on the interval $[0,t]$, we may replace it by a step function supported on the interval $[0, x_n]$ with the same norm, either by projecting f onto this interval (at a T stage) or by using Lemma 6 as a definition (at an S stage). We can then repeat this process, obtaining a function supported on $[0, x_{n-1}]$ and so on. After finitely many stages one obtains a function $\omega\chi_{[0,s]}$ for some $\omega, s \geqslant 0$. The norm of f is then defined to be $\omega\lambda(s)$. The resulting space we will call the ST-*space associated with* λ, denoted L_λ^t. (Note that our notation is not at all standard: L_ϕ usually stands for the Orlicz space associated with the Orlicz function ϕ. However, the notation is so convenient here that we have adopted it.)

During the rest of this section we will make a number of uses of the simple fact that if $X \leqslant Y$ then $T_\alpha X \leqslant T_\alpha Y$ and $S_\alpha X \leqslant S_\alpha Y$ (the function-space analogue of Lemma 2).

LEMMA 9. *Let $X \in \mathcal{G}_u$, let $Y_1, Y_2 \in \mathcal{G}_t$ be two ST-extensions of X and let λ_1 and λ_2 be the associated ST-functions. Then if $\lambda_1(x) \leqslant \lambda_2(x)$ for every $x \in [u,t]$, it follows that $Y_1 \leqslant Y_2$.*

PROOF. Let $w_1, \dots, w_N$ be the set of values, in increasing order, taken by either λ_1 or λ_2 when they are differentiable with derivative zero, or taken by λ_2 when it is maximal. Let s be maximal such that λ_1 and λ_2 are equal on the interval $[u,s]$. Then $\lambda_1(s) = \lambda_2(s) = w_i$, say. If $s < t$, then our aim is to replace λ_1 by a larger function λ_1', still dominated by λ_2, but now equal to it until they both take the value w_{i+1}. Write Z for the ST-extension of X corresponding to the restrictions of λ_1 and λ_2 to the interval $[u,s]$, and $U_1, \dots, U_l$ for the corresponding ST-sequence. Note that the next terms in the sequences of Y_1 and Y_2 must be of the form T_α and S_β, by the maximality of s. We must now consider various cases.

First, if $i + 1 = N$ and neither λ_1 nor λ_2 has zero right derivative when they take the value w_N, then w_N is the maximum of λ_2 and the ST-sequences of

Y_1 and Y_2 must be $U_1, \dots, U_l, T_\alpha$ and $U_1, \dots, U_l, S_\alpha$ respectively, (where α can be shown to be w_{i+1}/w_i). In this case, let $\lambda_1' = \lambda_2$, and observe that the only modification we have made to the ST-sequence corresponding to λ_1 is to replace the final T_α by S_α.

Otherwise, we know that at least one of λ_1 and λ_2 has zero derivative when it takes the value w_{i+1}. Suppose first that λ_1 has. Then the ST-sequence of the space Y_1 begins $U_1, \dots, U_l, T_{\alpha_1}, S_{\alpha_2}, T_{\alpha_3}$, with $\alpha_1, \alpha_2, \alpha_3 > 1$ (and again one can show that $\alpha_2 = w_{i+1}/w_i$). We modify this sequence by exchanging T_{α_1} with S_{α_2} and define λ_1' to be the correspondingly modified ST-function. It is not hard to show that λ_1' is still dominated by λ_2 and that λ_1' and λ_2 are equal until they both take the value w_{i+1}.

If on the other hand λ_2 has zero derivative when it takes the value w_{i+1}, then the ST-sequence of Y_2 begins $U_1, \dots, U_l, S_{\beta_1}, T_{\beta 2}$, with $\beta_1, \beta_2 > 1$. We now have three further sub-cases. If the ST-sequence of Y_1 is $U_1, \dots, U_l, T_{\alpha_1}$, then $\alpha_1 \geqslant \beta_1\beta_2$, so we can replace this ST-sequence by the equivalent sequence $U_1, \dots, U_l, T_{\beta_1}, T_\gamma$ where $\beta_1\gamma = \alpha_1$. Now modify the sequence by changing T_{β_1} into S_{β_1} and let λ_1' be the corresponding function. Then again λ_1' is dominated by λ_2 but equal to it until they both equal w_{i+1}.

If the ST-sequence of Y_1 begins $U_1, \dots, U_l, T_{\alpha_1}, S_{\alpha_2}$ and $\alpha_2 \geqslant \beta_1$, then we can replace it with the equivalent sequence $U_1, \dots, U_l, T_{\alpha_1}, S_{\beta_1}, S_\gamma$, where $\gamma = \alpha_2/\beta_1$. We modify this sequence by exchanging T_{α_1} and S_{β_1} and let λ_1' be the corresponding function. Once again, it has the required property.

Finally, if the ST-sequence of Y_1 begins $U_1, \dots, U_l, T_{\alpha_1}, S_{\alpha_2}$ and $\alpha_2 < \beta_1$, then the hypothesis of this case implies that $U_1, \dots, U_l, T_{\alpha_1}, S_{\alpha_2}$ is the whole sequence. In this case, we modify the sequence in two steps. First we exchange the T_{α_1} and the S_{α_2}. Then we replace the T_{α_1} with $S_{\gamma_1}, T_{\gamma_2}$ where $\gamma_1 = \beta_1/\alpha_2$ and $\gamma_2 = \alpha_1\alpha_2/\beta_1$.

In each of the cases above, λ_1' is obtained from λ_1 either by changing a T_α in its ST-sequence to an S_α, or by changing the order of a consecutive pair S_α, T_β so that after the change the S_α operation is performed first, (or in the final case doing both). By the definitions of the operations S_α and T_β, the simple fact mentioned just before this lemma, and Lemma 8, each such modification increases the corresponding norm. Hence, by induction on i, one can transform λ_1 into λ_2 by applying a finite sequence of changes to the corresponding ST-sequence, each of which increases the corresponding norm. It follows that $Y_1 \leqslant Y_2$ as stated. □

COROLLARY 10. *Let $X_1, X_2 \in \mathcal{G}_u$ and let λ_1 and λ_2 be two ST-functions on the interval $[u, t]$ with $\lambda_i(u) = \|\chi_{[0,u]}\|_{X_i}$ for $i = 1, 2$. Let Y_1 and Y_2 be the corresponding ST-extensions of X_1 and X_2. Suppose there is a constant $c > 0$ such that $X_1 \leqslant cX_2$, $\lambda_1(u) = c\lambda_2(u)$ and $\lambda_1(x) \leqslant c\lambda_2(x)$ for any $x \in [u, t]$. Then*

$$Y_1 \leqslant cY_2.$$

PROOF. Without loss of generality $c = 1$. By Lemma 9 and the analogue of Lemma 2 we see that $Y_1 \leqslant Z \leqslant Y_2$, where Z is the λ_2-extension of X_1. □

COROLLARY 11. *Let λ and μ be two ST-functions defined on the interval $(0, t]$ with $\lambda(s) \leqslant c\mu(s)$ for every s. Then $L_\lambda^t \leqslant cL_\mu^t$.*

PROOF. Let f be a decreasing step function with the first step ending at u. The norm of f in an ST-space does not depend on the behaviour of the ST function below u. Therefore for values less than u we may replace λ and μ by any other ST-functions. Choose functions $\lambda_1 \leqslant c\mu_1$ such that for some $u_1 \leqslant u$ we have that $\lambda_1(u_1) = c\mu_1(u_1)$ and λ_1 and μ_1 are linear below u_1. We may then apply Corollary 10 (where X_1 and X_2 will be multiples of $L_1^{u_1}$). □

LEMMA 12. *Let $s, t > 0$, let $1 < p < \infty$, let q be the conjugate index of p and let $\alpha > 1$. Then $S_\alpha(\alpha^{-1/q} L_p^s) \leqslant L_p^{\alpha s}$ and $L_p^{\beta t} \leqslant T_\beta(\beta^{1/p} L_p^t)$.*

PROOF. The second inequality states that, for any positive decreasing function f on the interval $[0, \beta t]$, we have the inequality

$$\int_0^{\beta t} (f(x))^p \, dx \leqslant \beta \int_0^t (f(x))^p \, dx,$$

which is obvious. To prove the first, observe that by the above inequality, with $t = \alpha^{-p/q} s$ and $\beta = \alpha^{p/q}$, we have

$$\alpha^{-1/q} L_p^s \leqslant T_\beta(L_p^t),$$

so, by the extension of Lemma 2, we have

$$S_\alpha(\alpha^{-1/q} L_p^s) \leqslant S_\alpha T_\beta(L_p^t).$$

Now $\beta = \alpha^{p/q} = \alpha^{p-1}$, so, by Lemma 7,

$$S_\alpha T_\beta(L_p^t) \leqslant L_p^{\alpha\beta t} = L_p^{\alpha s}.$$ □

We are now ready for the main theorem of this section.

THEOREM 13. *Let $0 < u < t$, let $c_1 \leqslant c \leqslant c_2$ and let λ be an ST-function on the interval $[u, t]$ such that $\lambda(u) = cu^{1/p}$ and $c_1 x^{1/p} \leqslant \lambda(x) \leqslant c_2 x^{1/p}$ for any $x \in [u, t]$. Let Y be the ST-extension of cL_p^u corresponding to the function λ. Then $c_1 L_p^t \leqslant Y \leqslant c_2 L_p^t$.*

PROOF. We shall prove only the left-hand inequality: the other is similar. Without loss of generality $c_1 = 1$. Define Z to be the space $S_{c^q} L_p^{c^{-q} u}$, the norm of which, by Lemma 12, is dominated by that of cL_p^u. Notice that $\|\chi_{[0,u]}\|_Z = cu^{1/p}$.

We now define an ST-sequence $S_{\alpha_1}, T_{\beta_1}, S_{\alpha_2}, T_{\beta_2}, \dots, S_{\alpha_k}, T_{\beta_k}$, letting μ be the function (which will be defined on the interval $[u, t]$) corresponding to the subsequence $T_{\beta_1}, S_{\alpha_2}, T_{\beta_2}, \dots, S_{\alpha_k}, T_{\beta_k}$. Set $\alpha_1 = c^q$ and $\beta_1 = \alpha_1^{p-1} = c^p$. In general, once we have defined $\alpha_1, \dots, \alpha_{i-1}$ and $\beta_1, \dots, \beta_{i-1}$, we let α_i be maximal such that the resulting part of the function μ is dominated by $\min\{\lambda, t^{1/p}\}$, and then set $\beta_i = \alpha_i^{p-1}$. This construction guarantees that, whenever $\mu'(a) = 0$, either there is some b such that $\lambda'(b) = 0$ and $\lambda(b) = \mu(a)$, or $\mu(a) = t^{1/p}$. (Informally, at each stage the logarithmic graph of μ rises with slope one until it hits a horizontal part of the logarithmic graph of λ, at which point it becomes horizontal until it rejoins the line $y = x/p$.) Since λ has only finitely many changes of slope, the graph of μ reaches a height of $t^{1/p}$ after only finitely many changes of direction, and after one further step (corresponding to T_{β_k}) the process stops.

Let X be the ST-extension of Z corresponding to the function μ. Then Lemma 7 implies (after an easy induction) that $L_p^t \leqslant X$. Since $Z \leqslant cL_p^u$ and $\mu \leqslant \lambda$ with $\mu(u) = \lambda(u)$, Corollary 10 implies that $X \leqslant Y$, completing the proof of the theorem. □

Corollary 14. *Let λ be an ST-function defined on the interval $(0,t]$ such that $c_1 s^{1/p} \leqslant \lambda(s) \leqslant c_2 s^{1/p}$ for every s. Then $c_1 L_p^t \leqslant L_\lambda^t \leqslant c_2 L_p^t$.*

Proof. Just as in the proof of Corollary 11, we can deduce this from Theorem 13 by considering step functions first. □

It remains to show that we can approximate ℓ_p^n by an ST-space in the sense of Section 2, whatever the value of p. Theorem 13 tells us that all we need to worry about is the norm of vectors of the form $\sum_1^k \boldsymbol{e}_i$.

Theorem 15. *Let $1 \leqslant p \leqslant \infty$. Then for any $n \in \mathbb{N}$, there exists an ST-space X such that $d(X, \ell_p^n) < 3/2$.*

Proof. By duality, we may assume that $1 \leqslant p \leqslant 2$. Suppose $X \in \mathcal{F}_n$, $X' \in \mathcal{G}_n$ and the embedding I_n is an isometry from X to its image in X'. It is not hard to see that I_{n+1} is an isometric embedding from $T(X)$ into $T_\alpha(X')$, and also from $S(X)$ into $S_\alpha(X')$, where $\alpha = 1+1/n$. By Theorem 13, then, it is enough to find an ST-sequence $U_1, U_2, \ldots, U_{n-1}$ such that each U_i is either $S_{1+1/i}$ or $T_{1+1/i}$ and the corresponding ST-function λ satisfies $x^{1/p} \leqslant \lambda(x) < (3/2)x^{1/p}$. Clearly we only need to check this inequality when x is an integer. Suppose we have chosen $U_1, \ldots, U_{k-1}$. Then if $\lambda(k) \geqslant (k+1)^{1/p}$, set $U_k = T_{1+1/k}$. Otherwise, let it be $S_{1+1/k}$. Then for each k, $\lambda(k)$ is either $\lambda(k-1)$ or it is less than $k^{1/p} \cdot k/(k-1)$. In the first case we are done, by induction. In the second, we are done, unless $k = 2$. But in this case, $\lambda(k) = 2 < (3/2) \cdot 2^{1/p}$. □

Note that Theorem 13 implies easily that if $k \in \mathbb{N}$ is even and $X \in \mathcal{F}_{2^k}$ is the space $T^{2^{k-1}}(S^{2^{k-2}}(\ldots(S^4(T^2(S(\mathbb{R}))))\ldots))$, then $d(X, \ell_2^{2^k}) = \sqrt{2}$. Indeed, X embeds isometrically into the space $T_2(S_2(\ldots(T_2(S_2(L_2^1)))\ldots)) \in \mathcal{G}_{2^k}$, and if λ is the ST-function associated with this space, then $\lambda(2^l) = 2^{l/2}$ if l is even, and $2^{(l+1)/2}$ if l is odd. It follows that $\sqrt{x} \leqslant \lambda(x) \leqslant \sqrt{2x}$ for every x, which, by Theorem 13, is enough to prove the above estimate. (As in Section 2, the lower bound on the distance follows from the fact that the distance is attained by the identity map.) This can also be proved directly from Lemma 7, using a duality argument similar to that of Theorem 1. In general, to approximate ℓ_p^n to within an absolute constant by an ST-space, one needs only about log n changes of direction of the corresponding ST-function λ, or log n terms in the corresponding ST-sequence. Our arguments show also that many other ST-sequences would have worked just as well in Theorem 1.

If one wishes for a better polytope approximation of ℓ_p^n, still with an easy geometrical description, then one method is to consider appropriate subspaces. For example, if $2m = nk$, $X = S(TS)^{m-1}(\mathbb{R})$ and Y is the n-dimensional subspace of X generated by the block basis $\boldsymbol{u}_i = \sum_{j=(i-1)k+1}^{ik} \boldsymbol{e}_j$, where $i = 1, 2, \ldots, n$, then Theorem 13 and a straightforward calculation show that $d(Y, \ell_2^n) \leqslant 1 + C/k$ for some absolute constant C.

5. Limits of ST-Spaces

We say that a function λ is a *growth function* if it is a strictly positive uniform limit of ST-functions. It is not hard to show that $\lambda : [u, t] \to \mathbb{R}$ is a growth function if and only if $\lambda(u) > 0$ and $\lambda(x) \leqslant \lambda(y) \leqslant (y/x)\lambda(x)$ whenever $u \leqslant x \leqslant y \leqslant t$, and also if and only if there is a space $X \in \mathcal{G}_t$ such that $\lambda(x) = \|\chi_{[0,x]}\|_X$ for every $u \leqslant x \leqslant t$. (A similar statement holds also for growth functions defined on the interval $(0, t]$.) It is an easy consequence of Lemma 10 that, given $X \in \mathcal{G}_u$ and a growth function λ such that $\lambda(u) = \|\chi_{[0,u]}\|_X$, the following normed space $Y \in \mathcal{G}_t$ is well defined. Pick any sequence $\lambda_1, \lambda_2, \ldots$ of ST-functions tending uniformly to λ with $\lambda_n(u) = \lambda(u)$ and let

$$\|f\|_Y = \lim_{n \to \infty} \|f\|_{Y_n}$$

where Y_n is the ST-extension of X corresponding to λ_n. We shall call the space Y the λ-*extension* of X. Given a uniform limit $\lambda : (0, t] \to \mathbb{R}$ of ST-functions λ_n, we can define a norm $\|\cdot\|_\lambda$ by again taking $\|f\|_\lambda$ to be the limit of the norms $\|f\|_{\lambda_n}$. This is the most natural definition, in the context of ST-spaces, of a norm associated with the growth function λ. We shall denote the completion of $(F[0, t], \|\cdot\|_\lambda)$ by L_λ^t. Later, we shall show that the space L_λ^t is in some ways more natural than the Orlicz space with the same growth function.

The main task of this section is to show how to calculate the norm of a function f in a λ-extension (and consequently in a space L_λ^t). This we do by giving a continuous analogue of the algorithm defined in Section 3. Thus, we are generalizing our results so far from growth functions with logarithmic gradient 0 or 1 to growth functions with logarithmic gradient in the interval $[0, 1]$. This generalization is of some interest for its own sake, but the immediate benefit is a second proof of Theorem 15 which avoids the use of Lemma 7.

For convenience, we shall assume that our growth function λ is differentiable with non-zero derivative, and we shall show how to calculate $\|f\|_\lambda$ when f is continuously differentiable with negative derivative and both f and f' are bounded away from 0 and ∞ (this includes right and left derivatives at 0 and 1). A function f with these properties we shall call *standard*. Using rearrangement-invariance and straightforward limiting arguments, one can deal with more general f and λ.

Given a standard function f, we would like, just as in Section 3, to replace f by a function of smaller support but with the same norm. Given $y \leqslant t$ large enough it will turn out that there is a unique $x \leqslant y$ such that if we define the function g_y by

$$g_y(s) = \begin{cases} f(s) & \text{if } 0 \leqslant s \leqslant x, \\ f(x) & \text{if } x \leqslant s \leqslant y, \\ 0 & \text{if } s > y, \end{cases}$$

then $\|g_y\|_\lambda = \|f\|_\lambda$. It is clear that x decreases as y decreases. (The differentiability of f applies provided x is greater than 0, which is the reason for the condition that y should be large enough.) Of much more interest is the fact that the dependence between x and $y = y(x)$ is given by the differential equation

$$y \frac{\lambda'(y)}{\lambda(y)} \frac{dy}{dx} = -\frac{f'(x)}{f(x)}(y - x), \tag{1}$$

which can be rewritten as

$$y \, d(\log \lambda(y)) = -(y - x) \, d(\log f(x)).$$

The main task of this section is to derive equation (1). However, let us first see why it gives a new proof of Theorem 15. Note first that when $x = 0$ (and y is maximal), the function g_y is simply $f(0)$ times the characteristic function of $[0, y]$, so that $\|f\|_\lambda = \|g_y\|_\lambda = f(0)\lambda(y(0))$. Hence, the differential equation gives us a means of calculating the norm $\|\cdot\|_\lambda$. Next, observe that if $\lambda(y) = y^{1/p}$, then the solution of equation (1) is

$$y = f(x)^{-p} \left(C - \int_0^x f(s)^p \, ds \right) + x$$

for some constant C. Since $y(t) = t$, we obtain that $C = \int_0^t f(s)^p \, ds$. Hence, $y(0) = f(0)^{-p} \int_0^t f(s)^p \, ds$. Thus $f(0)(y(0))^{1/p} = \|f\|_p$. But Corollary 10 (which did not use Lemma 7) can obviously be generalized to the same result for limits of ST-functions and the corresponding spaces. Since the function $t^{1/p}$ is such a function and it gives rise to the space L_p^t, we obtain Theorem 13 and hence Theorem 15.

To obtain equation (1), let $\varepsilon > 0$ and $0 < u < t$ and let $\alpha_1, \beta_1, \dots, \alpha_N, \beta_N$ be a sequence of real numbers greater than 1 with the following properties.

(i) $\lambda(\alpha_1\beta_1 \dots \alpha_k\beta_k u) = \beta_1 \dots \beta_k \lambda(u)$ for $k = 1, 2, \dots, N$.
(ii) $\alpha_k\beta_k \leqslant 1 + \varepsilon$ for $k = 1, 2, \dots, N$.
(iii) $\alpha_1\beta_1 \dots \alpha_N\beta_N = t/u$.

Let $X \in \mathcal{G}_u$ and let Y be the ST-extension $S_{\beta_N} T_{\alpha_N} \dots S_{\beta_1} T_{\alpha_1}(X)$ and notice that the growth function μ of Y has the following properties.

(a) $\mu(\alpha_1\beta_1\ldots\alpha_k\beta_k) = \lambda(\alpha_1\beta_1\ldots\alpha_k\beta_k)$ for $k = 1, 2, \ldots, N$.
(b) $(1+\varepsilon)^{-1}\lambda(s) \leqslant \mu(s) \leqslant \lambda(s)$ for $u \leqslant s \leqslant t$.

For each k, define $y_k = \alpha_1\beta_1\ldots\alpha_k\beta_k u$. Fix $k < N$, set $y = y_k$ and $\delta y = y_{k+1} - y_k$, and suppose that $x = x_k$ has been defined. Define a function g by the formula

$$g(s) = \begin{cases} f(s) & \text{if } 0 \leqslant s \leqslant x, \\ f(x) & \text{if } x \leqslant s \leqslant y, \\ 0 & \text{if } s > y, \end{cases}$$

and let δx be the unique number such that $\|h\|_\mu = \|g\|_\mu$, where

$$h(s) = \begin{cases} f(s) & \text{if } 0 \leqslant s \leqslant x + \delta x, \\ f(x+\delta x) & \text{if } x + \delta x \leqslant s \leqslant y + \delta y, \\ 0 & \text{if } s > y + \delta y. \end{cases}$$

Then define x_{k+1} to be $x + \delta x$. We shall now obtain an approximate equation relating δx and δy.

Setting $\alpha = \alpha_{k+1}$ and $\beta = \beta_{k+1}$, we know from the definition of the operation T_α that $\|g\|_\mu = \|g_1\|_\mu$, where

$$g_1(s) = \begin{cases} f(s) & \text{if } 0 \leqslant s \leqslant x, \\ f(x) & \text{if } x \leqslant s \leqslant \alpha y, \\ 0 & \text{if } s > \alpha y. \end{cases}$$

By Lemma 6 (which in this context could almost be regarded as the definition of S_β) we know also that $\int_0^t g_1(s)\,ds = \int_0^t h(s)\,ds$. That is,

$$\int_0^x f(s)\,ds + (\alpha y - x) f(x) = \int_0^{x+\delta x} f(s)\,ds + (\alpha\beta y - x - \delta x) f(x+\delta x),$$

which implies that

$$\int_x^{x+\delta x} (f(s) - f(x))\,ds + (x+\delta x)(f(x) - f(x+\delta x)) = \alpha y f(x) - \alpha\beta y f(x+\delta x).$$

Bearing in mind that $\delta x = O(\delta y)$ and that $(\alpha\beta - 1)y = \delta y$, which also implies that $\alpha - 1$ and $\beta - 1$ are $O(\delta y)$, we can simplify the above to

$$\frac{\beta - 1}{\alpha\beta - 1}\delta y + o(\delta y) = -\frac{f'(x)}{f(x)}(y - x)\delta x. \tag{2}$$

Finally, notice that

$$\lambda'(y) = \frac{\lambda(\alpha\beta y) - \lambda(y)}{(\alpha\beta - 1)y} + o(1) = \frac{\beta - 1}{\alpha\beta - 1}\frac{\lambda(y)}{y} + o(1),$$

so that, substituting into (2), we have the estimate

$$y\frac{\lambda'(y)}{\lambda(y)}\delta y + o(\delta y) = -\frac{f'(x)}{f(x)}(y - x)\delta x. \tag{3}$$

Letting ε, and hence δy, tend to zero, we see that in the limit as μ tends to λ, we do obtain equation (1) as claimed. Having obtained the equation for a λ-extension, it is easy to see that it is valid for the space L_λ^t as well.

We end this section with a small remark. It is easy to prove that the dual of L_λ^t is L_μ^t where $\mu(s) = s/\lambda(s)$, first when λ is an ST-function, and then, on taking limits, for an arbitrary growth function. (This statement is not quite accurate since for example the dual of L_∞^t is not L_1^t. What we mean is that the norm of a measurable function in the dual of L_λ^t is its L_μ^t-norm.) The proof comes straight from the definition of the operations S_α and T_β. Similarly, it is trivial that these spaces are normed spaces. Therefore we have an argument which makes the inequalities of Hölder and Minkowski in some sense "obvious" and "geometrical". Unfortunately, working out the details is more complicated than the usual proofs of those inequalities!

6. Two Results about ST-Spaces

We shall be concerned with two natural questions in this section. First, what is the relationship, if anything, between ST-spaces and Orlicz spaces? Second, what is the result of interpolating between two ST-spaces?

For the first question, suppose X is an Orlicz function space restricted to the interval $[0, t]$ with the norm given by

$$\|f\|_X = \inf \left\{ \mu > 0 : \int_0^t \phi(|f(s)|/\mu)\, ds \leqslant 1 \right\}$$

where ϕ is an Orlicz function. For this space we have $\|\chi_{[0,s]}\|_X = (\phi^{-1}(s^{-1}))^{-1}$. For $0 < u < t$ let X_u be the restriction to $[0, u]$ of X. If for $u \leqslant s \leqslant t$ we set $\lambda(s) = (\phi^{-1}(s^{-1}))^{-1}$, it is natural to ask whether the (completion of the) λ-extension of X_u is X? We shall show that it is, isometrically, if and only if, for some p, $\phi(s) = s^p$ and therefore $\lambda(s) = s^{1/p}$. In other words, ST-spaces and Orlicz spaces intersect only in the L_p-spaces.

Our proof of this is slightly indirect. Suppose ϕ is an Orlicz function, X is the corresponding Orlicz space and X_u and λ are given as above. Suppose moreover that the identity is an isometry from X to the λ-extension of X_u. Let f be any standard function. For $y < t$ sufficiently close to t we replace f by a function g_y with the properties we had in the last section. That is, for some $0 \leqslant x \leqslant y$,

$$g_y(s) = \begin{cases} f(s) & \text{if } 0 \leqslant s \leqslant x, \\ f(x) & \text{if } x \leqslant s \leqslant y, \\ 0 & \text{if } s > y, \end{cases}$$

and $\|g_y\|_\lambda = \|f\|_\lambda$. It is easy to see that, for each y, g_y is unique. Suppose that $\|f\|_X = 1$. Using the definition of the norm in X and the isometry assumption one readily obtains that

$$\int_0^x \phi(f(s))\ ds + (y-x)\phi(f(x)) = \int_0^t \phi(f(s))\ ds \tag{4}$$

If we differentiate this equation with respect to x and rearrange, we obtain that

$$\frac{\phi(f(x))}{\phi'(f(x))f(x)}\frac{dy}{dx} = -\frac{f'(x)}{f(x)}(y-x).$$

On the other hand, if we substitute $\lambda(y) = (\phi^{-1}(y^{-1}))^{-1}$ into the differential equation (1), we obtain the equation

$$\frac{1}{y\phi^{-1}(y^{-1})\phi'(\phi^{-1}(y^{-1}))}\frac{dy}{dx} = -\frac{f'(x)}{f(x)}(y-x).$$

It follows from the uniqueness of g_y that $\phi(f(x))/\phi'(f(x))f(x)$ does not depend on $f(x)$. In other words, the function $t \mapsto \phi(t)/t\phi'(t)$ is a constant function. Solving this equation gives $\phi(t) = Ct^p$ for some constants C and p. We have proved the next theorem.

Theorem 16. *Let $X \in \mathcal{G}_t$ be the restriction of an Orlicz function space on $[0,t]$ to the interval $[0,t]$ and let Y be the completion of the λ-extension of X_u, where λ and X_u are as defined above. Then X is isometric to Y under the identity map if and only if, for some constants $C > 0$ and $1 \leqslant p \leqslant \infty$, $\|f\|_X = \|f\|_Y = C\|f\|_p$.*

It would be much more interesting to find out when ST-spaces are *isomorphic* to Orlicz spaces. There seems to be a reasonable chance that they always are, in which case they could be regarded as the "correct" renorming of Orlicz spaces.

The next result shows that ST-spaces interpolate in the way one would expect. Since we prove an isometric result, we must use complex interpolation and therefore complex scalars. One can either define the norm of any vector to be the norm of its modulus, or follow the original approach making obvious modifications. These give the same result. Also, we shall make use of the fact (related to the remark at the end of the previous section) that, if $X \in \mathcal{G}_u$, and λ is a growth function, then the dual of the λ-extension of X is the μ-extension of X^*, where $\mu(s) = s/\lambda(s)$. (Again, this statement should be interpreted somewhat loosely.) For the basic facts and notation to do with interpolation, see [BL].

Theorem 17. *Let $X \in \mathcal{G}_u$, let λ and μ be growth functions and let X_λ and X_μ be the λ- and μ-extensions of X respectively. Given $0 < \theta < 1$, let $\nu = \lambda^\theta\mu^{1-\theta}$ and let X_ν be the ν-extension of X. Then*

$$(X_\lambda, X_\mu)_{[\theta]} = X_\nu.$$

Proof. The proof of this is very similar to the standard proof that L_p-spaces interpolate in the way one would expect. Let f be a standard function with

$\|f\|_{X_\nu} = 1$ and let $y = y(x)$ be defined by the differential equation

$$y\frac{\nu(y)}{\nu(y)}\frac{dy}{dx} = -\frac{f'(x)}{f(x)}(y-x)$$

and the initial condition $y(t) = t$. Let $D \subset \mathbb{C}$ be the set $\{z \in \mathbb{C} : 0 \leqslant \mathrm{Re}(z) \leqslant 1\}$. Fixing y, we may now define, for every $z \in D$, a function g_z by the differential equation

$$y\left((1-z)\frac{\lambda'(y)}{\lambda(y)} + z\frac{\mu'(y)}{\mu(y)}\right)\frac{dy}{dx} = -\frac{g_z'(x)}{g_z(x)}(y-x)$$

and the initial condition

$$g_z(0) = \lambda(y(0))^{-(1-z)}\mu(y(0))^{-z}.$$

It is easy to check that

$$g_z(x) = g_0(x)^{1-z}g_1(x)^z. \tag{3}$$

Now set $\tilde{f}(z,x) = \exp(\varepsilon z^2 - \varepsilon\theta^2)g_z(x)$. Thus $\tilde{f} : D \times [0,t] \to \mathbb{C}$. We have the properties of $\tilde{f}$ necessary to estimate $\|f\|_{(X_\lambda,X_\mu)_{[\theta]}}$. First, for each x it is clear that $\tilde{f}(z,x)$ is analytic in z on the interior of D. Second, $\|\tilde{f}(ir,\cdot)\|_{X_\lambda}$ and $\|\tilde{f}(1+ir,\cdot)\|_{X_\mu}$ both tend to zero as $|r|$ tends to infinity. Moreover, we have $\|\tilde{f}(ir,\cdot)\|_{X_\lambda} \leqslant 1$ and $\|\tilde{f}(1+ir,\cdot)\|_{X_\mu} \leqslant \exp(\varepsilon)$ for every $r \in \mathbb{R}$. It follows that $\|f\|_{(X_\lambda,X_\mu)_{[\theta]}} \leqslant \exp(\varepsilon)$. Since $\varepsilon > 0$ was arbitrary, we have shown that $\|f\|_{(X_\lambda,X_\mu)_{[\theta]}} \leqslant \|f\|_{X_\nu}$ for any function f.

Conversely, suppose that $\|f\|_{(X_\lambda,X_\mu)_{[\theta]}} = 1$. This tells us that for each $\varepsilon > 0$ there exists a function $\tilde{f}$ with the above properties. We also know that

$$\|f\|_{X_\nu} = \sup\left\{|\langle f,h\rangle| : h \text{ standard}, \|h\|_{X_{\nu_1}} = 1\right\}$$

where $\nu_1(s) = s/\nu(s)$.

Given a standard function h with $\|h\|_{\nu_1} \leqslant 1$, let $\tilde{h} : D \times [0,t] \to \mathbb{C}$ be given by the method used to construct $\tilde{f}$ from f, replacing all the spaces in that construction by their duals. Then set

$$F(z) = \int_0^t \tilde{f}(z,x)\tilde{h}(z,x)\,dx$$

for every $z \in D$. Then F is analytic on the interior of D, continuous on D, and $F(ir) \leqslant 1$ and $F(1+ir) \leqslant \exp(2\varepsilon)$ for every $r \in \mathbb{R}$. By the Hadamard three-line theorem (see [BL]) we obtain that

$$|\langle f,h\rangle| \leqslant |F(\theta)| \leqslant \exp(2\varepsilon).$$

Since $\varepsilon > 0$ was arbitrary, we have $\|f\|_{X_\nu} \leqslant 1$. ☐

It is not hard to deduce from Theorem 17 that, with the same notation, we have also $(L^t_\lambda, L^t_\mu)_{[\theta]} = L^t_\nu$.

7. Suggestions for Further Research

There are many questions one can ask arising from the results of the previous sections. Some of this section is extremely speculative.

(**A**) The most urgent question is whether ST-spaces are renormings of Orlicz spaces or something quite different. This question has been mentioned earlier in the paper so we shall not say much more about it. Whatever the answer, one can ask whether ST-spaces always contain some ℓ_p almost isometrically. We have an argument, whose details are yet to be checked, that they do; it uses the similar result of Lindenstrauss and Tzafriri for Orlicz spaces, for which see [LT].

(**B**) It is likely that the map taking a growth function λ to the rearrangement-invariant function space $L_\lambda(\mathbb{R})$ with that growth function can be characterized amongst all such maps. Here are some properties that might be involved:

(i) $t^{1/p} \mapsto L_p(\mathbb{R})$.
(ii) $t/\lambda(t) \mapsto (L_\lambda(\mathbb{R}))^*$.
(iii) $a\lambda \leqslant \mu$ implies $aL_\lambda(\mathbb{R}) \leqslant L_\mu(\mathbb{R})$.
(iv) $(L_\lambda^t, L_\mu^t)_{[\theta]} = L^t_{\lambda^\theta \mu^{1-\theta}}$.

Are these enough? Perhaps (iii) needs to be a little more detailed. For example, one also has:

(v) If $a\lambda(s) \leqslant \mu(s)$ for $s \leqslant t$ and the support of f is contained in the interval $[0, t]$, then $a\|f\|_\lambda \leqslant \|f\|_\mu$.
(vi) If f^* is constant on $[0, u]$ and $a\lambda(s) \leqslant \mu(s)$ for $s \geqslant u$, then $a\|f\|_\lambda \leqslant \|f\|_\mu$.

(**C**) Recall the remark at the end of Section 4, that only $\log n$ changes between S and T are needed to approximate ℓ_p^n to within a constant. This is clearly best possible. Does it cause the unit ball of the resulting ST-space to have interesting extremal properties of a more geometrical nature amongst polytopes approximating the ball of ℓ_p^n? Unchecked calculations suggest that the number of vertices and faces are both at most exponential (again for $1 < p < \infty$ this is necessary). If they are correct, then we have constructed efficient approximating polytopes in a very explicit way. Perhaps this might have algorithmic uses. It would be interesting to have good estimates for the number of facets of each dimension of the polytope arising from a given ST-sequence.

The restriction of L_λ^n to the subspace generated by the functions $\chi_{[i-1,i)}$ seems to have a polytope as its unit ball when λ is an ST-function (rather than a more general growth function). One can ask similar questions about these polytopes.

(**D**) It is interesting to define ST-spaces more geometrically, especially if we return to the idea of trying to construct "generic" polytopes. This can be done as follows. Given an affine subspace Y of $\mathbb{R}^n$ not containing zero, define the *canonical extension* of Y to a hyperplane to be the sum of Y and the orthogonal complement of the linear subspace generated by Y. Let the side of this hyperplane containing zero be the *canonical half-space* associated with Y.

Call a set of points $X = \{x_1, \ldots, x_m\}$ in $\mathbb{R}^n$ *compatible* if the canonical half-spaces associated with the (zero-dimensional) affine subspaces $\{x_i\}$ all contain all of X. Suppose we have a compatible set X in general position (for simplicity only; this condition is not really needed) and let Σ be the simplicial complex of all subsets $A \subset X$ such that $\operatorname{conv} A$ is a facet of $\operatorname{conv} X$. If $A \in \Sigma$, denote by $C(A)$ the cone generated by A. Let $C_r = C_r(X)$ be the union of all $C(A)$ such that $|A| = r+1$ (so that $C(A)$ is r-dimensional).

Given a subset $K \subset C_r$ such that $\operatorname{conv} K$ is a polytope with 0 in its interior and such that $K = C_r \cap \operatorname{conv} K$, we define $SK \subset C_{r+1}$ to be the intersection of C_{r+1} with $\operatorname{conv} K$, and we define TK to be the intersection of C_{r+1} with the intersection of all canonical half-spaces associated with r-facets of $\operatorname{conv} K$ that lie entirely in C_r. This definition agrees with the previous one if X is the set of points $\pm e_i$. It is possible to make precise the sense in which the S and T operations are dual to each other.

Several questions arise immediately. If one takes a suitably well-distributed set of points in the n-sphere and applies an ST-sequence which would have produced an approximation to ℓ_p^n if applied to $\{\pm e_i\}$, then what does one get, or at least what properties can one expect the resulting polytope (or normed space in the symmetric case) to have? (A possible definition of "suitably well-distributed" is that the identity on $\mathbb{R}^n$ can be expressed in the form $y \mapsto \sum c_i \langle x_i, y\rangle x_i$ for positive constants c_i and points x_i in the set.) What happens if one starts with a regular simplex? In this case, does alternating between S and T produce an approximation to a sphere? If so, what can one say about "simplex ℓ_p^n" for $p \neq 2$? (One might, for example, expect that the unit balls of simplex ℓ_p^n and simplex ℓ_q^n were equivalent via a negative multiple of the identity.) Is it possible to define a space given a more general growth function and a compatible set of starting points?

(**E**) One can regard ST-spaces as the result of a kind of interpolation between L_1 and L_∞. We give an indication of how this is done. Given a space $X \in G_t$ and a decreasing function f supported on $[0, \alpha t]$, we calculate its norm in $S_\alpha(X)$ by finding a function g of the form $f \vee \gamma$ restricted to $[0, t]$, where γ is minimal such that $\|g\|_1 = \|f\|_1$. Its norm in $T_\alpha(X)$ can be described in exactly the same way, except that now $\|g\|_\infty = \|f\|_\infty$ (so that $\gamma = 0$). Given a rearrangement-invariant function space V, we could define $V_\alpha(X)$ in the same way. The norm of f in $V_\alpha(X)$ is the norm in X of the restriction of $f \vee \gamma$ to $[0, t]$, where now γ is chosen so that the V-norm of the two functions is the same.

If we have two RI-spaces V and W, and an ST-sequence with corresponding growth function λ, we can replace each S_α by a V_α and each T_α by a W_α. Denote the resulting space by $(V, W)_{[\lambda]}$. Then $(L_1, L_\infty)_{[\lambda]} = L_\lambda$. (We have extended to the whole of $\mathbb{R}_+$—this presents no problems.) We can now take limits as before and define $(V, W)_{[\lambda]}$ for arbitrary growth functions λ.

Some questions that arise out of this definition are the following. What is the relationship between this interpolation method and the complex interpolation method, when $\lambda(t) = t^{1/p}$? It is not hard to show that $(L_\lambda, L_\mu)_{[\nu]} = L_\xi$, where ξ is defined by the differential equation

$$\frac{\xi'(x)}{\xi(x)} = \frac{\nu'(x)}{\nu(x)}\frac{\lambda'(x)}{\lambda(x)} + \Big(1 - \frac{\nu'(x)}{\nu(x)}\Big)\frac{\mu'(x)}{\mu(x)}\,,$$

so they agree with complex interpolation when V and W are limits of ST-spaces.

Does this method give rise to an interpolation theorem, and is the resulting constant 1? This would be interesting as it is a real method rather than a complex one. Can the method be generalized to interpolation between arbitrary spaces? (Probably it is not hard to generalize at least as far as lattices, but even this is not a triviality.)

(**F**) It seems very likely that it is possible to generalize many of the results of this paper to operator spaces. Do appropriate operator-space versions give operator ℓ_p-spaces, as defined by Pisier? Does this give a means of defining operator Orlicz spaces? Of course the answer to this last question can only be yes if ST-spaces and Orlicz spaces coincide isomorphically.

Acknowledgement

This paper is a revised and expanded version of a chapter in my PhD thesis, which contained all the main results. The thesis was written under the supervision of Dr B. Bollobás at Cambridge University, to whom I am very grateful. I would also like to thank the referee, whose long and detailed report, full of valuable suggestions, reached me less than two weeks after I submitted the paper. As a result, the presentation has been significantly improved and I have been saved from several errors.

References

[BL] J. Bergh and J. Löfström, Interpolation Spaces, an Introduction, Springer, Berlin, 1976.

[G] E. D. Gluskin, *The diameter of the Minkowski compactum is approximately equal to n*, Funct. Anal. Appl. **15** (1981), 72–73.

[LT] J. Lindenstrauss and L. Tzafriri, Classical Banach spaces I: sequence spaces, Springer, Berlin, 1977.

W. Timothy Gowers
Department of Pure Mathematics and Mathematical Statistics
Cambridge University
16 Mill Lane
Cambridge CB2 1SB
United Kingdom
W.T.Gowers@dpmms.cam.ac.uk