Lecture Notes in Mathematics

Edited by A. Dold and B. Eckmann

1200

Vitali D. Milman
Gideon Schechtman

Asymptotic Theory of Finite Dimensional Normed Spaces

With an Appendix by M. Gromov

Springer-Verlag

Lecture Notes in Mathematics

Edited by A. Dold and B. Eckmann

1200

Vitali D. Milman
Gideon Schechtman

Asymptotic Theory of Finite Dimensional Normed Spaces

With an Appendix by M. Gromov
Isoperimetric Inequalities in Riemannian Manifolds

Springer-Verlag
Berlin Heidelberg New York London Paris Tokyo

Mathematics Subject Classification (1980): 46B20, 52A20, 60F10

ISBN 3-540-16769-2 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-16769-2 Springer-Verlag New York Berlin Heidelberg

Printing and binding: Beltz Offsetdruck, Hemsbach/Bergstr.
2146/3140-543210

INTRODUCTION

This book deals with the geometrical structure of finite dimensional normed spaces, as the dimension grows to infinity. This is a a part of what came to be known as the Local Theory of Banach Spaces (this name was derived from the fact that in its final stages, this theory dealt mainly with relating the structure of infinite dimensional Banach spaces to the structure of their lattice of finite dimensional subspaces).

Our purpose in this book is to introduce the reader to some of the results, problems, and mainly methods developed in the Local Theory, in the last few years. This by no means is a complete survey of this wide area. Some of the main topics we do not discuss here are mentioned in the Notes and Remarks section. Several books appeared recently or are going to appear shortly, which cover much of the material not covered in this book. Among these are Pisier's [Pis6] where factorization theorems related to Grothendieck's theorem are extensively discussed, and Tomczak-Jaegermann's [T-J1] where operator ideals and distances between finite dimensional normed spaces are studied in detail. Another related book is Pietch's [Pie].

The first major result of the Local Theory is Dvoretzky's Theorem [Dv] of 1960. Dvoretzky proved that every real normed space of finite dimension, say n, contains a $(1+\varepsilon)$-isomorphic copy of the k-dimensional euclidean space ℓ_2^k, for $k = k(\varepsilon, n)$ which increases to ∞ with n (see Chapter 5 for the precise statement). Dvoretzky's original proof was very complicated and understood only by a few people. In 1970 Milman [M1] gave a different proof which exploited a certain property of the Haar measure on high dimensional homogeneous spaces, a property which is now called the concentration phenomenon: Let (X, ρ, μ) be a compact metric space (X, ρ) with a Borel probability measure μ. The concentration function $\alpha(X, \varepsilon)$, $\varepsilon > 0$, is defined by

$$\alpha(X,\varepsilon) = 1 - inf\{\mu(A_\varepsilon);\ \mu(A) \geq \frac{1}{2},\ A \subseteq X\ Borel\}$$

where

$$A_\varepsilon = \{x \in X;\ \rho(x, A) \leq \varepsilon\}\ .$$

It turns out that for some very natural families of spaces, $\alpha(X, \varepsilon)$ is extremely small. For example, it follows from Levy's isoperimetric inequality that for the Euclidean n-sphere S^n, with the geodesic distance ρ and the normalized rotational invariant measure μ,

$$\alpha(S^n, \varepsilon) \leq \sqrt{\frac{\pi}{8}} exp(-\varepsilon^2 n/2)\ .$$

It follows from this inequality (see Chapter 2) that any nice real function on S^n must be very close to being a constant on all but a very small set (the exceptional set being of measure of order smaller than $exp(-\varepsilon^2 n/2)$). This last property is what is called the concentration phenomenon. It has proved to be extremely useful in the study of finite dimensional normed spaces.

Going back to the concentration function, we define a family (X_n, ρ_n, μ_n) of metric probability spaces to be a Levy family if $\alpha(X_n, \varepsilon_n \; diam \; X_n) \underset{n\to\infty}{\longrightarrow} 0$ ($diam \; X_n$ is the diameter of X_n). Chapter 6 below contains a lot of examples of such natural families. Many of these examples have deep applications in the Local Theory. It is usually quite a difficult task to establish that a certain family is a Levy family, the methods are different from one example to the other and come from diverse areas (including methods from differential geometry, estimation of eigenvalues of the Laplacian, large deviation inequalities for martingales, isoperimetric inequalities). Levy families, the concentration phenomenon and their applications to the asymptotic theory of normed spaces are the main topics of the first part of this book. We have already mentioned one application, namely Dvoretzky's Theorem. In the same direction we deal with estimation of the dimension of euclidean subspaces of various large families of normed spaces, that is, with the evaluation of the function $k(\varepsilon, n)$ mentioned above when restricted to some (wide) families of normed spaces. (This study originated in [M1] and [F.L.M.].) Here is an example: There exists a function $c(\varepsilon) > 0$, $\varepsilon > 0$, such that if $X_n = (\mathbb{R}^n, \|\cdot\|)$ is a family of normed spaces, then either for some $\alpha > 0$ and any $\varepsilon > 0$ and n, X_n contains a $(1+\varepsilon)$-isomorphic copy of ℓ_2^k with $k = [c(\varepsilon)n^\alpha]$, or for any integer k and any $\varepsilon > 0$, there is an n such that X_n contains a $(1+\varepsilon)$-isomorphic copy of ℓ_∞^k. (The proof of this result uses, besides the concentration phenomenon, also the notions of type and cotype introduced below.)

We also deal, in the first part, with packing high dimensional ℓ_p^k, $1 < p < 2$, spaces into ℓ_1^n (Chapter 7) as well as packing spaces with special structure (unconditional or symmetric bases) into general normed spaces (Chapter 10).

The second part of the book revolves around the notions of type and cotype and the relation of these notions to the geometry of normed spaces.

For $1 \le p \le 2 \le q \le \infty$ the type p constant (resp. cotype q constant) of X denoted $T_p(X)$ $(C_q(X))$ is the smallest constant in the inequality

$$\left(\underset{\varepsilon_i=\pm 1}{\text{Ave}} \|\sum_{i=1}^k \varepsilon_i x_i\|^2\right)^{1/2} \le T \left(\sum_{i=1}^k \|x_i\|^p\right)^{1/p}$$

$$\left(\quad \left(\sum_{i=1}^k \|x_i\|^q\right)^{1/q} \le C \left(Ave\|\sum_{i=1}^k \varepsilon_i x_i\|^2\right)^{1/2} \quad \right)$$

for all k and $x_1, \ldots, x_k \in X$.

These notions were introduced by Hoffmann-Jorgensen [H-J] for the study of limit theorems for vector valued random variables, and were studied extensively by Maurey and Pisier ([M.P.] in particular) in connection with the geometry of normed spaces. They have proved to be a very important tool in the Local Theory. In particular, Krivine, Maurey and Pisier showed that for $p_X = sup\{p;\ T_p(X) < \infty\}$ (resp. $q_X = inf\{q;\ C_q(X) < \infty\}$), $\ell^n_{p_X}$ (resp. $\ell^n_{q_X}$) are $(1+\varepsilon)$-isomorphic to subspaces of X for all n and $\varepsilon > 0$ (here X is infinite dimensional; there is also a corresponding statement for finite dimensional spaces). We present a proof (somewhat different from previous proofs) of this theorem in Chapters 12 and 13. (Chapter 11 deals with some infinite dimensional combinatorial methods needed in the sequel.)

Chapter 14 is devoted to the work of Pisier [Pis1] estimating the norm of one specific projection called the Rademacher projection. This is closely related to the relation between the type p constant of X and the cotype q constant of X^* $(\frac{1}{p}+\frac{1}{q}=1)$. It also has applications to finding well complemented euclidean sections in normed spaces. These applications due to Figiel and Tomczak-Jaegermann [F.T.] are discussed in Chapter 15.

The book also contains five appendices, the first of which is written by M. Gromov and gives an introduction to the theory of isoperimetric inequalities on riemannian manifolds. It is written in a way understandable to the non-expert (in Differential Geometry). This appendix contains also results which were not published elsewhere (in particular – the Gromov-Levy isoperimetric inequality). We are indebted to M. Gromov for this excellent addition to our book.

CONTENTS

Appendices

PART I: THE CONCENTRATION OF MEASURE PHENOMENON IN THE THEORY OF NORMED SPACES

1. PRELIMINARIES.

In this chapter we present some preliminary material on invariant measures needed in later chapters. Not all the details of the proofs are given.

1.1. Let (M, ρ) be a compact metric space and let G be a group whose members act as isometries on M, i.e. for $g \in G$, $t, s \in M$, $\rho(gt, gs) = \rho(t, s)$.

We begin by introducing the Haar measure. The proof we give is due apparently to W. Maak (see [Do]).

THEOREM: *There exists a regular measure μ on the Borel subsets of M which is invariant under the action of members of G, i.e., $\mu(A) = \mu(gA)$ for all $A \subseteq M$, $g \in G$. Alternatively, $\int f(t)d\mu(t) = \int f(gt)d\mu(t)$ for all $g \in G$ and $f \in C(M)$. ($C(M)$ is the linear space of all real continuous functions on M).*

PROOF: For each $\varepsilon > 0$ let N_ε be a minimal ε-net in M, i.e., $\cup_{t \in N_\varepsilon} B(t, \varepsilon) = M$ and $n_\varepsilon = |N_\varepsilon|$, the cardinality of N_ε, is minimal among all sets with this property. ($B(t, \varepsilon) = \{s \in M; \rho(t, s) \leq \varepsilon\}$).

For $f \in C(M)$ define

$$\mu_\varepsilon(f) = n_\varepsilon^{-1} \sum_{t \in N_\varepsilon} f(t) .$$

Since $\{\mu_\varepsilon\}_{\varepsilon > 0}$ is a uniformly bounded set of linear functionals on $C(M)$, it follows (see e.g. [D.S.]) that, for some sequence $\varepsilon_i \underset{i \to \infty}{\longrightarrow} 0$,

$$\mu_{\varepsilon_i}(f) \underset{i \to \infty}{\longrightarrow} \mu(f)$$

for all $f \in C(M)$ where $\mu(\cdot)$ is a linear positive functional on $C(M)$ with $\mu(1) = 1$. That is, μ is given by a regular Borel probability measure (see [D.S.] again).

Next we want to show that the measure μ is uniquely determined, i.e., if μ'_ε is defined using a different minimal ε-net, then for the same sequence ε_i, $\mu'_{\varepsilon_i}(f) \to \mu(f)$. If N'_ε is another minimal ε-net in M then we claim: there exists a one to one and onto map $\varphi: N_\varepsilon \to N'_\varepsilon$ with $\rho(t, \varphi(t)) \leq 2\varepsilon$ for all $t \in N_\varepsilon$.

To show this we use a combinatorial result known as the "marriage Theorem" (see e.g.[Do]). Let us say that $t \in N_\varepsilon$ and $s \in N'_\varepsilon$ are acquainted if $B(t,\varepsilon) \cap B(s,\varepsilon) \neq \emptyset$. Then the members of any subset K of N_ε are collectively acquainted with a subset L of N'_ε of at least as many elements as K. Indeed, let $L = \{s \in N'_\varepsilon; B(s,\varepsilon) \cap (\cup_{t\in K} B(t,\varepsilon)) \neq \emptyset\}$ then $|L| \geq |K|$ since otherwise $L \cup (N_\varepsilon \backslash K)$ is an ε-net with less then $|N_\varepsilon|$ elements. The marriage Theorem states that in such a situation (where every subset of N_ε is jointly acquainted with at least as many elements of N'_ε) there is a one to one map $\varphi: N_\varepsilon \to N'_\varepsilon$ with t and $\varphi(t)$ acquainted for all $t \in N_\varepsilon$. This translates back into $B(t,\varepsilon) \cap B(\varphi(t),\varepsilon) \neq \emptyset$, i.e., $\rho(t,\varphi(t)) \leq 2\varepsilon$. It follows that, if μ'_ε is defined using N'_ε in an analogous way to μ_ε, then

$$|\mu_\varepsilon(f) - \mu'_\varepsilon(f)| \leq \frac{1}{n_\varepsilon} \sum_{t\in N_\varepsilon} |f(t) - f(\varphi(t))| \leq \omega(2\varepsilon)$$

where $\omega(\varepsilon) = sup\{|f(t) - f(s)|; \rho(t,s) \leq \varepsilon\}$ is the modulus of continuity of f. Thus, $\lim_{i\to\infty} \mu'_{\varepsilon_i}(f)$ exists and is equal to $\mu(f)$.

It remains to show that μ is invariant under G. Let $g \in G$ and let $N'_\varepsilon = (gt)_{t\in N_\varepsilon}$. Then N'_ε is a minimal ε-net and

$$\begin{aligned}\mu(f \circ g) &= \lim_{i\to\infty} \frac{1}{n_{\varepsilon_i}} \sum_{t\in N_{\varepsilon_i}} f(gt) = \lim_{i\to\infty} \frac{1}{n_{\varepsilon_i}} \sum_{s\in N'_{\varepsilon_i}} f(s) \\ &= \lim_{i\to\infty} \mu'_{\varepsilon_i}(f) = \mu(f) \ .\end{aligned}$$

□

1.2. If the action of G on M is transitive, i.e. for all $t,s \in M$ there exists a $g \in G$ such that $gt = s$, then M is called an *homogeneous space* of G.

Fix a $t \in M$ and let

$$G_o = \{g \in G; gt = t\}$$

then G_o is a subgroup of G (called an *isotropic subgroup*) and $M = G/G_o$ where $s \in M$ is identified with the equivalence class gG_o for g such that $gt = s$.

To illustrate the definition we give a simple example : Fix an inner product $(\cdot,\cdot)$ on $\mathbb{R}^n$, let $|x| = (x,x)^{1/2}$ for $x \in \mathbb{R}^n$ and let $G = O_n$ be the orthogonal group on $(\mathbb{R}^n, \{\cdot,\cdot\})$. We identify O_n with the set of n tuples $(e_1,\dots e_n)$ of orthonormal vectors (Fix one orthonormal basis $(e_1^o,\dots,e_n^o)$ then any orthogonal operator A uniquely determines another such n-tuple: $(Ae_1^o,\dots,Ae_n^o)$). Let $M = S^{n-1} = \{x \in \mathbb{R}^n; |x| = 1\}$ and $\varphi: O_n \to S^{n-1}$ be defined by $\varphi(e_1,\dots,e_n) = e_1$. Then clearly for any $t \in S^{n-1}, \varphi^{-1}(t)$ can be identified with O_{n-1}. So $S^{n-1} = O_n/O_{n-1}$.

1.3. THEOREM: *If (M,ρ) is a compact metric homogeneous space of the group G then the measure of Theorem 1.1. is unique up to a constant.*

PROOF: Define a semi metric on G by $d(g,h) = sup_{t\in M}\rho(gt,ht)$. Identifying elements whose distance apart is zero, we get a group H which still acts as isometries on M and also on itself. (there are two ways in which H acts on itself - we choose multiplication on the right $hh' = h' \cdot h$ where " $\cdot$ " is the multiplication in the group). One checks that H is compact (actually in all our applications $G = H$ will be given as a compact group). Let μ on M and ν on H be measures invariant under the action of G. Then, for all $f \in C(M)$,

$$\nu(1)\mu(f) = \int_G \int_M f(gt)d\mu(t)d\nu(g) = \int_M \int_G f(gt)d\nu(g)d\mu(t).$$

By transitivity of G on M and invariance of ν, the inner integral on the right depends on f but not on t. Call it $\overline{\nu}(f)$. Then

$$\nu(1)\mu(f) = \overline{\nu}(f)\mu(1).$$

So that if μ' is another invariant measure on M then

$$\mu(f)\mu'(1) = \mu'(f)\mu(1)$$

□

1.4. REMARKS: **a.** The proof shows also that any right invariant normalized measure on a compact metric group G is equal to any normalized left invariant measure.

b. It is easily checked that the unique normalized invariant measure on G is also invertable invariant i.e. $\int_G f(t)d\mu(t) = \int_G f(t^{-1})d\mu(t)$.

In what follows μ will denote the normalized Haar measure on the space in question so that it may appear twice in the same formula denoting measures on different spaces.

1.5. We pass now to several examples of homogeneous spaces of the group O_n of all $n \times n$ real orthogonal matrices.

EXAMPLES: **a.** $S^{n-1} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n; \sum_{i=1}^n x_i^2 = 1\}$ with either the euclidean or geodesic metric is easily seen to be equivalent to O_n/O_{n-1} as was discussed above (1.2.).

b. *The Stiefel manifolds.* For $1 \leq k \leq n$

$$W_{n,k} = \{e = (e_1, \ldots, e_k)); e_i \in \mathbb{R}^n, (e_i, e_j) = \delta_{ij}, 1 \leq i, j \leq k\}$$

with the metric $\rho(e, f) = (\sum_{i=1}^k d(e_i, f_i)^2)^{\frac{1}{2}}$, d being either the euclidean or the geodesic metric. Note that $W_{n,n} = O_n, W_{n,1} = S^{n-1}$ and $W_{n,n-1} = SO_n = \{T \in O_n;\, det\ T = 1\}$. In general $W_{n,k}$ may be identified with O_n/O_{n-k} via the map $\varphi: O_n \to W_{n,k}$, $\varphi(e_1, \ldots, e_n) = (e_1 \ldots e_k)$.

c. The *Grassman manifolds* $G_{n,k}$, $1 \leq k \leq n$, consists of all k dimensional subspaces of $\mathbb{R}^n$ with the metric being the Hausdorff distance between the unit balls of the two subspaces

$$\rho(\xi, \varsigma) = \sup_{x \in S^{n-1} \cap \varsigma} \rho(x, S^{n-1} \cap \xi) .$$

The equivalence $G_{n,k} = O_n/(O_k \times O_{n-k})$ is again easily verified.

d. If G is any group with invariant metric ρ and G_o is a subgroup we may define a metric d on $M = G/G_o$ by

$$d(t,s) = inf\{\rho(g,h); \varphi(g) = t, \varphi(h) = s\}$$

where $\varphi: G \to M$ is the quotient map. In this way M becomes an homogeneous space of G. Note that in all previous examples the metric given on the homogeneous space of O_n is equivalent, up to a universal constant (not depending on n), to the metric given here.

1.6. The uniqueness of the normalized Haar measure allows us to deduce several interesting consequences.

a. The first remark is that for any $A \subseteq S^{n-1}$ and $x_o \in S^{n-1}$.

$$\mu\{T \in O_n;\ Tx_o \in A\} = \mu(A).$$

b. Next we give two identities. Fix $1 \le k \le n$, for $\xi \in G_{n,k}$ we denote $S(\xi) = S^{n-1} \cap \xi$ the $(k-1)$-dimensional sphere of ξ. Then

$$\int_{S^{n-1}} f d\mu = \int_{G_{n,k}} \int_{S(\xi)} f(t) d\mu_\xi(t) d\mu(\xi)$$

for all $f \in C(S^{n-1})$ where μ_ξ is the normalized Haar measure on $S(\xi)$ (by our convention μ on the left is the normalized Haar measure on S^{n-1} and on the right on $G_{n,k}$.

We identify $\mathbb{R}^{2n}$ with $\mathbb{C}^n$ (by introducing a complex structure in one of the possible ways). For each k we denote the collection of complex k-dimensional subspaces of $\mathbb{C}^n$ by $\mathbb{C}\, Gn, k$ and the unit sphere of any $\xi \in \mathbb{C}\, G_{n,k}$ by $\mathbb{C}\, S(\xi)$ (which can be identified with S^{2k-1}). $\mathbb{C}\, G_{n,k}$ is again an homogeneous space and we get an identity similar to the previous one

$$\int_{S^{2n-1}} f d\mu = \int_{\mathbb{C}\, G_{n,k}} \int_{\mathbb{C}\, S(\xi)} f(t) d\mu_\xi(t) d\mu(\xi).$$

Note that here one integrates on a much smaller space, $\mathbb{C}\, Gn, k$, than the one in the first identity which, adjusted to the dimensions here, would be $G_{2n,2k}$.

2. THE ISOPERIMETRIC INEQUALITY ON S^{n-1} AND SOME CONSEQUENCES

2.1. We begin with the statement of the classical isoperimetric inequality on the sphere. A proof is given in Appendix I. A simple proof of a version of Corollary 2.2. (which is sufficient for the applications here) is given in Appendix V.

For a set A in a metric space (M,ρ) and $\varepsilon > 0$ we denote $A_\varepsilon = \{t; \rho(t,A) \le \varepsilon\}$. In what follows we use the geodesic metric on S^{n-1}.

THEOREM: *For each* $0 < a < 1$ *and* $\varepsilon > 0$, $min\{\mu(A_\varepsilon);\ A \subseteq S^{n-1},\ \mu(A) = a\}$ *exists and is attained on* A_o *- a cap of suitable measure (i.e.,* $A_o = B(x_o, r)$ *for any* $x_o \in S^{n-1}$ *and* r *such that* $\mu(B(r)) = a$*, where* $B(r) = B(x_o,r) = \{x; \rho(x,x_o) \le r\}$*).*

Using this for the case $a = 1/2$ (in which case A_o is half a sphere) we get:

2.2. COROLLARY: *if* $A \subseteq S^{n+1}$ *with* $\mu(A) \ge 1/2$ *then* $\mu(A_\varepsilon) \ge 1 - \sqrt{\pi/8}e^{-\varepsilon^2 n/2}$

PROOF: By theorem 2.1 it is enough to evaluate $\mu(B(\pi/2+\varepsilon))$. Note that $cos^n\theta$ is proportional to the n-volume of S_θ - the set of points on S^{n+1} which are of distance θ from $B(1/2)$ (this is an n-dimensional sphere of radius cos θ). It follows easily (draw a picture) that

$$h(\varepsilon,n) = \mu(B(\pi/2+\varepsilon)) = \int_{-\pi/2}^{\varepsilon} \cos^n \theta d\theta \Big/ \int_{-\pi/2}^{\pi/2} \cos^n \theta d\theta \ .$$

Let $I_n = \int_0^{\pi/2} \cos^n \theta d\theta$ and use the change of variables $\theta = \tau/\sqrt{n}$ and the inequality $\cos t \le e^{-t^2/2}$, $0 \le t \le \pi/2$, to get

$$\begin{aligned}1 - h(\varepsilon,n) &= \int_\varepsilon^{\pi/2} \cos^n(\theta) d\theta/2I_n = (1/\sqrt{n}) \int_{\varepsilon\sqrt{n}}^{(\pi/2)\sqrt{n}} \cos^n(\tau/\sqrt{n}) d\tau/2I_n \\ &\le (1/\sqrt{n}) \int_{\varepsilon\sqrt{n}}^{(\pi/2)\sqrt{n}} e^{-\tau^2/2} d\tau/2I_n \le (1/\sqrt{n}) e^{-\varepsilon^2 n/2} \int_0^{(\pi/2-\varepsilon)\sqrt{n}} e^{-t^2/2} dt/2I_n \\ &\le (1/\sqrt{n}) e^{-\varepsilon^2 n/2} \int_0^\infty e^{-t^2/2} dt/2I_n = \frac{e^{-\varepsilon^2 n/2}\sqrt{\pi/2}}{2\sqrt{n}I_n} \ .\end{aligned}$$

This computation (with $\varepsilon = 0$) also gives $\sqrt{n}\ I_n \le \sqrt{\pi/2}$. To evaluate I_n from below notice that integration by parts gives $I_k = ((k-1)/k) I_{k-2}$ which implies $\sqrt{k}\ I_k \ge \sqrt{k-2}\ I_{k-2}$. Therefore,

$$\sqrt{n}\ I_n \ge min(I_1, \sqrt{2}\ I_2) = min(1, \sqrt{2}\ \pi/4) = 1$$

and

$$1 - h(\varepsilon, n) \le \sqrt{\pi/8}\; e^{-\varepsilon^2 n/2} .$$

□

REMARK: It can actually be checked that $\sqrt{n}\; I_n \underset{n\to\infty}{\longrightarrow} \sqrt{\pi/2}$.

2.3. The next corollary and theorem are crucial for the applications of the isoperimetric inequality to Banach spaces.

For a continuous function f on S^{n+1} we denote by $\omega_f(\varepsilon)$ its modulus of continuity, $\omega_f(\varepsilon) = \sup\{|f(x) - f(y)|; \rho(x,y) \le \varepsilon\}$. We denote by M_f the *median* (also called the *Levy mean)* of f, i.e., M_f is a number such that both

$$\mu\{x \in S^{n+1}; f(x) \le M_f\} \ge 1/2 \textit{ and } \mu\{x \in S^{n+1}; f(x) \ge M_f\} \ge 1/2.$$

COROLLARY (Levy's lemma): *Let* $f \in C(S^{n+1})$ *and let* $A = \{x; f(x) = M_f\}$ *then* $\mu(A_\varepsilon) \ge 1 - \sqrt{\pi/2}\; e^{-\varepsilon^2 n/2}$.

PROOF: Note that $A_\varepsilon = (f \le M_f)_\varepsilon \cap (f \ge M_f)_\varepsilon$ and use 2.2.

□

Notice that the values of f on A_ε are very close to M_f. Indeed, if ε is such that $\omega_f(\varepsilon) \le \delta$ then $|f(x) - M_f| \le \delta$ on A_ε. So the content of the previous corollary is that *a well behaved function is "almost" a constant on "almost" all the space.* This phenomenon of *concentration of measure* around one value of the function will appear over and over again in these notes.

2.4. In the next theorem we trade off the set of large measure on which the function is almost constant with a set with linear structure - a subspace of large dimension. The symbol $[\cdot]$ denotes the integer part function.

THEOREM: *For* $\varepsilon, \theta > 0$ *and an integer* n *let* $k(\varepsilon,\theta,n) = [\varepsilon^2 n/(2 log\ 4/\theta)]$. *Let* $f \in C(S^{n+1})$ *then, for all* $\varepsilon, \theta > 0$, *there exists a subspace* $E \subseteq \mathbb{R}^{n+2}$ *with* $dim\ E = k \ge k(\varepsilon,\theta,n)$ *and a* $\theta - net$ N in $S(E) = S^{n+1} \cap E$ *such that*

(i) $|f(x) - M_f| \le \omega_f(\varepsilon)$ *for all* $x \in N$

and

(ii) $|f(x) - M_f| \le \omega_f(\varepsilon) + \omega_f(\theta)$ *for all* $x \in E \cap S^{n+1}$.

(ii) follows from *(i)*. The proof of *(i)* consists of the following two lemmas. As in 2.3. $A = \{x; f(x) = M_f\}$.

2.5. LEMMA: *For any* $N \subseteq S^{n+1}$ *with* $|N| < \sqrt{\pi/2}\; e^{\varepsilon^2 n/2}$ *there exists a* $T \in O_{n+2}$ *such that* $TN \subseteq A_\varepsilon$. *Consequently, for all* $x \in TN$, $|f(x) - M_f| \le \omega_f(\varepsilon)$.

PROOF: The lemma follows immediately from 1.6.a: for each $x \in S^{n+1}$ $\mu\{T \in O_{n+2};\ Tx \in A_\varepsilon\} = \mu(A_\varepsilon) \ge 1 - \sqrt{\pi/2}\; e^{-\varepsilon^2 n/2}$. Therefore,

$$\mu\{T \in O_{n+2};\ \ Tx \in A_\varepsilon\ \ for\ every\ x \in N\} \geq 1 - |N|\sqrt{\pi/2}\ e^{-\varepsilon^2 n/2} > 0\,.$$

□

2.6. For later use we state the next lemma in a more general framework.

LEMMA: *For every normed space X with dim $X = k$ there exists a $\theta - net$ N in $S(X) = \{x \in X; \|x\| = 1\}$ with $|N| \leq (1 + 2/\theta)^k \leq e^{k\ log 3/\theta}$.*

PROOF: Let $\{x_i\}_{i=1}^n$ be a maximal set in $S(X)$ with the property that $\|x_i - x_j\| \geq \theta$ for $i \neq j$. Then $\{x_i\}_{i=1}^n$ is an $\theta - net$. The open balls $B(x_i, \theta/2)$ are pairwise disjoint and are all contained in $B(0, 1+\theta/2)$. Comparing the volume of $B(0, 1+\theta/2)$ with that of $\cup_{i=1}^n B(x_i, \theta/2)$ we get

$$n \cdot (\theta/2)^k \leq (1 + \theta/2)^k$$

or

$$n \leq (1 + 2/\theta)^k$$

□

2.7. REMARK: An inspection of the proof shows that, taking a bit smaller k in Theorem 2.4. one can get the conclusion for most subspaces of dimension k : For $k \leq [\varepsilon^2 n/(4 log\ 4/\theta)]$ *the measure of the set $E_k \subseteq G_{n+2,k}$ of all k dimensional subspace for which the conclusions of Theorem 2.4 hold satisfies $\mu(E_k) \geq 1 - \sqrt{\pi/2}\ e^{-\varepsilon^2 n/4}$.*

2.8. We now indicate briefly a different way to prove Theorem 2.4.

First we estimate the measure of a cup in S^{k-1} from below

$$\mu(B(\theta)) = \frac{\int_o^\theta \sin^{k-2} t dt}{2\ I_{k-2}} \geq \frac{\int_{\theta/2}^\theta \sin^{k-2} t dt}{2\ I_{k-2}} \geq \frac{\theta \sin^{k-2} \theta/2}{4\ I_{k-2}}.$$

Next we use the identity 1.6.b to get that there exists a k-dimensional subspace E such that

$$\mu(E \cap A_\varepsilon) \geq 1 - \sqrt{\pi/2}\ e^{-\varepsilon^2 n/2}\,.$$

If

$$\frac{\theta \sin^{k-2} \theta/2}{4\ I_{k-2}} + 1 - \sqrt{\pi/2}\ e^{-\varepsilon^2 n/2} \geq 1 \qquad (*)$$

then any ball of radius θ in E intersects A_ε and then for any $x \in E$, $|f(x) - M_f| \leq \omega(\varepsilon + \theta)$. Now use (*) to get an estimate on k.

2.9. REMARK: Note that if we want to prove a theorem similar to 2.4. in which the conclusion holds for two functions simultaneously (on the same subspace) we loose very little in the estimate on the dimension. Actually, by Remark 2.7., one can find a subspace E

with dim $E = k \geq [\varepsilon^2 n/(4 log\ 4/\theta)]$ for which the conclusion holds for exp $([\varepsilon^2 n/(4 log\ 4/\theta)])$ functions simultaneously.

2.10. We conclude this chapter with a few exercises.

a. Let $A \subseteq S^{n-1}$. Then either A_ε or $(A^c)_\varepsilon$ contains $S^{n-1} \cap E$ for some $k(\varepsilon/2, \varepsilon/2, n)$ - dimensional subspace E. ($k(\varepsilon, \theta, n)$ was defined in theorem 2.4.).

b. Fix $k \leq k(\varepsilon/2, \varepsilon/2, n)$, let $A \subseteq S^{n-1}$ be such that for all k-dimensional subspaces E, $A \cap E \neq \emptyset$. Then there exists a k-dimensional subspace E, with $S^{n-1} \cap E \subseteq A_{2\varepsilon}$.

HINT: Consider the function $\rho = \rho(x, A)$ and prove that the median M_ρ is at most ε.

c. For every $A \subseteq S^{n-1}$ we define

$$B_{A,k} = \{E \in G_{n,k}; E \cap A \neq \emptyset\}$$

and

$$I_{A,k} = \{E \in G_{n,k}; S(E) \subseteq A\} .$$

Fix $\varepsilon > 0$ and $k \leq [\varepsilon^2 n/10\ log \varepsilon^{-1}]$. Prove: if $\mu(B_{A,k}) > 2e^{-\varepsilon^2 n/4}$ then $\mu(I_{A_{4\varepsilon},k}) \geq 1 - 2e^{-\varepsilon^2 n/4}$.

HINT: First use Remark 2.7 to show $\mu(A_{2\varepsilon}) \geq 1/2$.

3. FINITE DIMENSIONAL NORMED SPACES, PRELIMINARIES

3.1. Let X, Y be two n-dimensional normed spaces. The *Banach-Mazur distance* between them is defined as

$$d(X,Y) = inf\{\|T\| \cdot \|T^{-1}\|;\ T: X \to Y\ isomorphism\}.$$

Obviously $d(X,Y) \geq 1$ and $d(X,Y) = 1$ if and only if X and Y are isometric. If $d(X,Y) \leq \lambda$ we say that X and Y are $\lambda-$ *isomorphic*. The notion of the distance also has a geometrical interpretation. If $d(X,Y)$ is small then in some sense the two unit balls $B(X) = \{x \in X; \|x\| \leq 1\}$ and $B(Y) = \{y \in Y; \|y\| \leq 1\}$ are close one to the other. More precisely there is a linear transformation φ such that

$$B(X) \subseteq \varphi(B(Y)) \subseteq d(X,Y)B(X).$$

The Banach-Mazur distance satisfies a multiplicative triangle inequality $(d(X,Z) \leq d(X,Y) \cdot d(Y,Z))$. Also $d(X^*, Y^*) = d(X,Y)$ for all X and Y where * denotes the dual space.

In the next few chapters we will consider the space $\mathbb{R}^n$ with two norms on it. One is a general norm $\|\cdot\|$, the other will always be an euclidean norm $|x| = (x,x)^{1/2}$ induced by some inner product $(\cdot,\cdot)$. We denote $D = \{x \in \mathbb{R}^n; |x| \leq 1\}$ and for $E \subseteq \mathbb{R}^n$, $S(E) = \{x \in E; |x| = 1\}$.

Let a, b be such that

3.1.1. $$a^{-1}|x| \leq \|x\| \leq b|x| \text{ for all } x \in \mathbb{R}^n\ .$$

We may define the norm, dual to $\|\cdot\|$ relative to $(\cdot,\cdot)$, by

$$\|x\|^* = \sup\left\{\frac{|(x,y)|}{\|y\|};\ y \in \mathbb{R}^n \setminus \{0\}\right\}.$$

We get easily that

3.1.2. $$b^{-1}|x| \leq \|x\|^* \leq a|x| \quad \text{for all } x \in \mathbb{R}^n\ .$$

Indeed, $|x|^2 = (x,x) \leq \|x\|\ \|x\|^* \leq b|x|\ \|x\|^*$, which gives the left side inequality. On the other hand for all $x, y \in \mathbb{R}^n \setminus \{0\}$

$$\frac{|(x,y)|}{\|y\|} \leq \frac{|x| \cdot |y|}{\|y\|} \leq a|x|$$

so that $\|x\|^* = \sup\left\{\frac{|(x,y)|}{\|y\|}\ ;\ y \neq 0\right\} \leq a|x|$.

3.2. The following two theorems deal with the ellipsoid of maximal volume inscribed in the unit ball of a normed space. We recall that a centrally symmetric ellipsoid in $\mathbb{R}^n$ is the body