March 13, 1985

We have been looking at the periodicity map

$$[0,\pi] \times \mathcal{I}_k(V) \longrightarrow \mathcal{I}_{k-1}(V)$$
$$(\theta, J) \longmapsto (\cos\theta)\gamma^k + (\sin\theta)J$$

Here V is a C_k-module with inner product and $\mathcal{I}_k(V)$ = space of s.a. involution J anti-commuting with $\gamma^1, \ldots, \gamma^k$. What I want to do now is to find the character forms on $\mathcal{I}_k(V)$ and to establish their compatibility under periodicity.

Look at $k=1$, and write ε for γ^1. Then $V = V^+ \oplus V^-$ where

$$\varepsilon = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad J = \begin{pmatrix} 0 & g^{-1} \\ g & 0 \end{pmatrix} \qquad g \text{ unitary}$$

and

$$F_\theta = \cos\theta\,\varepsilon + \sin\theta\, J$$
$$= \begin{pmatrix} \cos\theta & \sin\theta\cdot g^{-1} \\ \sin\theta\cdot g & \cos\theta \end{pmatrix}$$

$$e_\theta = \frac{1+F_\theta}{2} = \begin{pmatrix} \frac{1+\cos\theta}{2} & \frac{\sin\theta}{2}g^{-1} \\ \frac{\sin\theta}{2}g & \frac{+1-\cos\theta}{2} \end{pmatrix}$$

is the orthogonal projector onto the subspace

$$\operatorname{Im}\begin{pmatrix} 1 \\ xg \end{pmatrix} \qquad x = \frac{\sin\theta}{1+\cos\theta} = \frac{2\sin\theta/2\,\cos\theta/2}{2(\cos\theta/2)^2}$$
$$= \tan(\theta/2).$$

Thus $\theta/2$ is the angle in the graph construction.

The character forms on the Grassmannian are

$$\frac{1}{2^{2k+1}\,k!}\ \mathrm{tr}\ F(dF)^{2k}$$

because in terms of the projector $e = \frac{F+1}{2}$ we know the curvature is $e(de)^2$ and the k-th character form is $\frac{1}{k!}\ \mathrm{tr}\ e(de)^{2k}$.

Now we want to compute the pull back of the kth character form under the map $\theta \mapsto F_\theta$ followed by integrating θ from 0 to π.

$$F_\theta = \cos\theta\ \varepsilon + \sin\theta\ J$$

$$dF_\theta = d\theta\left[-\sin\theta\ \varepsilon + \cos\theta\ J\right] + \sin\theta\ dJ$$

These two terms in dF_θ commute because dJ anti-commutes with $d\theta$, ε, J. Thus

$$i(\partial_\theta)\, F_\theta\, (dF_\theta)^{2k} = F_\theta\ 2k\,(-\sin\theta\,\varepsilon + \cos\theta\, J)\,(\sin\theta\ dJ)^{2k-1}$$

$$= 2k\,(\cos\theta\ \varepsilon + \sin\theta\ J)(-\sin\theta\ \varepsilon + \cos\theta\ J)\,(\sin\theta)^{2k-1}(dJ)^{2k-1}$$

$$= 2k\,(-\sin^2\theta\ J\varepsilon + \cos^2\theta\ \varepsilon J)(\quad)(\quad)$$

$$= 2k\,(\sin\theta)^{2k-1}\ \varepsilon J\,(dJ)^{2k-1}$$

$t = \sin^2\theta$
$dt = 2\sin\theta\cos\theta\, d\theta$

Need

$$2k\int_0^\pi (\sin\theta)^{2k-1}\,d\theta = 2k\cdot\int_0^{\pi/2}(\sin\theta)^{2k-2}(\cos\theta)^{-1}\,2\sin\theta\cos\theta\,d\theta$$

$$= 2k\cdot\int_0^1 t^{k-1}(1-t)^{-1/2}\,dt = 2k\,\frac{\Gamma(k)\Gamma(\frac12)}{\Gamma(k+\frac12)}$$

$$= 2k\,\frac{(k-1)!\ \cancel{\sqrt{\pi}}}{\cancel{\sqrt{\pi}}\ \frac12\ \frac32 \cdots \frac{2k-1}{2}} = \frac{2^{k+1}\,k!}{1\cdot 3\cdots(2k-1)}$$

So the kth character form pulled back and integrated gives

$$\frac{1}{2^{2k+1}k!}\,\frac{2^{k+1}k!}{1\cdot3\cdots(2k-1)} = \frac{k!}{2^k k!\,1\cdot3\cdots(2k-1)} = \frac{k!}{(2k)!}$$

$$= \frac{1}{2}\,\frac{(k-1)!}{(2k-1)!}$$

times $\operatorname{tr}\left(\varepsilon J(dJ)^{2k-1}\right)$. Thus the odd character form of degree $2k-1$ is

$$\boxed{\frac{1}{2}\,\frac{(k-1)!}{(2k-1)!}\,\operatorname{tr}\left(\varepsilon J\,(dJ)^{2k-1}\right)}$$

This checks for if $J=\begin{pmatrix}0 & g^{-1}\\ g & 0\end{pmatrix}$, then

$$\operatorname{tr}\left(\varepsilon J(dJ)^{2k-1}\right) = \operatorname{tr}\left(\begin{pmatrix}1&0\\0&-1\end{pmatrix}\begin{pmatrix}0 & g^{-1}\\ g & 0\end{pmatrix}\begin{pmatrix}0 & dg^{-1}\\ dg & 0\end{pmatrix}^{2k-1}\right)$$

$$= \operatorname{tr}\left\{g^{-1}dg\,(dg^{-1}dg)^{k-1} - g\,dg^{-1}(dg\,dg^{-1})^{k-1}\right\}$$

$$= (-1)^{k-1}\,2\,\operatorname{tr}\,(g^{-1}dg)^{2k-1}$$

Next take $k=2$. Here the family is

$$F_\theta = (\cos\theta)\gamma^2 + (\sin\theta)J$$

and the differential form to be pulled back is

$$\frac{1}{2}\,\frac{k!}{(2k+1)!}\,\operatorname{tr}\left(\gamma^{1}\cdot F\,(dF)^{2k+1}\right)$$

So we calculate as before

$$dF_\theta = d\theta\left[-\sin\theta\,\gamma^2 + \cos\theta\, J\right] + \sin\theta\, dJ$$

$$F_\theta (dF_\theta)^{2k+1} = (2k+1)\, d\theta\,(\cos\theta\,\gamma^2 + \sin\theta\, J)(-\sin\theta\,\gamma^2 + \cos\theta\, J) \times (\sin\theta\, dJ)^{2k}$$

$$= (2k+1)\, d\theta\, (\sin\theta)^{2k}\, \gamma^2 J (dJ)^{2k}$$

so we get

$$\underbrace{(2k+1)\,\frac{1}{2}\,\frac{k!}{(2k+1)!}\underbrace{\int_0^\pi (\sin\theta)^{2k} d\theta}_{\frac{\Gamma(k+\frac{1}{2})\Gamma(\frac{1}{2})}{\Gamma(k+1)} = \frac{\pi\,\frac{1}{2}\cdot\frac{3}{2}\cdots\frac{2k-1}{2}}{k!}}}_{\frac{1}{2}\,\frac{k!}{\underbrace{\cancel{(2k)!}}_{2^k k!}}\,\frac{\pi\,\cancel{(1\cdot 3\cdots 2k-1)}}{2^k k!}} \quad tr\left(\gamma^1\gamma^2 J (dJ)^{2k}\right)$$

$$= \frac{1}{2^{2k+1}k!}\,\pi\, tr\left(\gamma^1\gamma^2 J (dJ)^{2k}\right)$$

But then recall that $J = \begin{pmatrix} 0 & -iF \\ iF & 0 \end{pmatrix} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} F$

$\gamma^1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ $\gamma^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, so

$$\gamma^1\gamma^2 \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} = iI$$

and so the end answer is

$$\boxed{(2i\pi)\,\frac{1}{2^{2k+1}\,k!}\, tr\left(F (dF)^{2k}\right)}$$

Let me now try to summarize what we have just done. We first described the basic Bott periodicity maps. The model we use for the k-th representing space for the K-theory is as follows. We take a large C_k-module with inner product V and let $\mathcal{J}_k(V)$ = space of self-adjoint involutions on V which anti-commute with $\gamma^1, \ldots, \gamma^k$. Then one has the periodicity map

$$[0,\pi] \times \mathcal{J}_k(V) \longrightarrow \mathcal{J}_{k-1}(V)$$

$$(\theta, J) \longmapsto F_\theta = (\cos\theta)\gamma^k + (\sin\theta) J$$

Ex. $k=1$. $\gamma^1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ on $V = V^+ \oplus V^-$

$J = \begin{pmatrix} 0 & g^{-1} \\ g & 0 \end{pmatrix}$

$$F_\theta = \begin{pmatrix} \cos\theta & \sin\theta\, g^{-1} \\ \sin\theta\, g & -\cos\theta \end{pmatrix}, \quad e = \frac{1+F_\theta}{2} = \begin{pmatrix} \frac{1+\cos\theta}{2} & \frac{\sin\theta}{2} g^{-1} \\ \frac{\sin\theta}{2} g & \frac{1-\cos\theta}{2} \end{pmatrix}$$

is the projector onto $\operatorname{Im}\begin{pmatrix} 1 \\ xg \end{pmatrix}$, where

$$x = \frac{\sin\theta}{1+\cos\theta} = \tan(\theta/2).$$

So the periodicity map is the graph construction.

Ex $k=2$: $\gamma^1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ $\gamma^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ on $V = W \oplus W$

and $J = \begin{pmatrix} 0 & -iF \\ iF & 0 \end{pmatrix}$ with $F^2 = 1$ on W. Then

$$F_\theta = \begin{pmatrix} 0 & \cos\theta - i\sin\theta\, F \\ \cos\theta + i\sin\theta\, F & 0 \end{pmatrix}$$

corresponds to the path of unitary operator $\cos\theta + i\sin\theta\, F$ joining

I to $-I$.

The second thing we did was to describe the Chern character forms on $\mathcal{J}_k(V)$, at least for $k=0,1,2$ and check that they are compatible with the periodicity maps.

For $k=0$ the j-th character form is

$$\frac{1}{2^{2j+1}\,j!}\ \mathrm{tr}\ F(dF)^{2j}$$

For $k=1$ the character form of degree $2j-1$ is is

$$\frac{1}{2}\,\frac{(j-1)!}{(2j-1)!}\ \mathrm{tr}\left(\gamma^1 F(dF)^{2j-1}\right).$$

For $k=2$ I get something like

$$\frac{1}{2^{2j+1}\,j!}\ \mathrm{tr}\left(\gamma^1\gamma^2 F(dF)^{2j}\right)$$

up to a factor of $2i$.

Summary:

I have been thinking about Bott periodicity. On the first level it amounts to explicit homotopy equivalences

$$U \sim \Omega BU \qquad\qquad BU \sim \Omega U.$$

I constructed maps

1) $\qquad U \longrightarrow \Omega BU \qquad\qquad BU \longrightarrow \Omega U$

using the graph map in the former and the

Cayley transform in the latter.

Then I unify the two constructions using Clifford algebras. Let V be an ungraded C_k-module with inner product, and let $\mathcal{I}_k(V)$ be the space of self-adjoint involutions on V which anti-commute with $\gamma^1, \dots, \gamma^k$. This space is alternately a Grassmannian or unitary group depending on the parity of k because of the algebraic periodicity of the Clifford algebras. The maps 1) are part of maps

$$\mathcal{I}_k(V) \longrightarrow \Omega\, \mathcal{I}_{k-1}(V)$$

So far I have just constructed maps, but I have not explained why they are homotopy equivalences. I have not yet introduced the operators which are [redacted] essential [redacted] for the proof of periodicity.

March 19, 1985

M manifold, E real vector bundle over M, X section of E transversal to the zero section, Z the subm. of M where X vanishes. Over M one has a complex of vector bundles

$$\longrightarrow \Lambda^2 E^* \xrightarrow{X\lrcorner} \Lambda^1 E^* \xrightarrow{X\lrcorner} \Lambda^0 E^*$$

which is acyclic off Z. Claim the complex of top. vector spaces (Frechet spaces)

$$\longrightarrow \Gamma(M,\Lambda^2E^*) \longrightarrow \Gamma(M,\Lambda^1E^*) \longrightarrow \Gamma(M) \longrightarrow \Gamma(Z) \to 0$$

is split exact.

I believe this is true and will try to work out a proof.

The first thing to try to prove is that it is local on M. Suppose we have an open covering $\{U_\alpha\}$ of M and contracting homotopies h_α over U_α. I can assume by passing to a refinement that the covering has a subordinate partition of unity $\sum \rho_\alpha = 1$. Then

$$h = \sum_\alpha \rho_\alpha h_\alpha \iota_\alpha \qquad \begin{array}{l}\iota_\alpha \text{ restriction}\\ \text{from } M \text{ to } U_\alpha\end{array}$$

will be a contracting homotopy over M.

What is going on here is we have a complex $K(U)$:

$$\longrightarrow \Gamma(U,\Lambda^1E^*) \longrightarrow \Gamma(U,\Lambda^0E^*) \longrightarrow \Gamma(U\cap Z) \longrightarrow 0$$

and an embedding and projection

$$K(M) \underset{i}{\overset{\rho}{\rightleftarrows}} \prod_\alpha K(U_\alpha).$$

March 15, 1985

Suppose given $F = \begin{pmatrix} 0 & P^{-1} \\ P & 0 \end{pmatrix}$ on $H = H^+ \oplus H^-$ where $H^\pm$ are A-modules and F is q-summable. Then we have two ways of constructing cyclic cocycles attached to F, φ_n n even $n >> q$.

1) Connes' approach: The operator F determines a map $K_0 A \longrightarrow \mathbb{Z}$ as follows. Given an idempotent matrix e of size r, one forms $P^{\oplus r} : (H^+)^r \longrightarrow (H^-)^r$ and reduces by e to obtain $eP^{\oplus r}e : e(H^+)^r \longrightarrow e(H^-)^r$. This is Fredholm and it has an index. One has

$$\mathrm{Ind}(ePe) = (-1)^k \, tr(\varepsilon\, e[F,e]^{2k}) = \frac{(-1)^k}{2} tr(\varepsilon F [F,e]^{2k+1})$$

provided the traces make sense.

This motivates the expression

$$\begin{aligned} 2\varphi_n(a_0, \dots, a_n) &= tr(\varepsilon F [F,a_0] \cdots [F,a_n]) \\ &= tr(\varepsilon (\underset{1}{F^2} a_0 - F a_0 F)[F,a_1] \cdots [F,a_n]) \\ &= 2\, tr(\varepsilon\, a_0 [F,a_1] \cdots [F,a_n]) \end{aligned}$$

Better:

$$\boxed{\begin{aligned} \varphi_n(a_0, \dots, a_n) &= tr(\varepsilon\, a_0 [F,a_1] \cdots [F,a_n]) \\ &= \tfrac{1}{2} tr(\varepsilon F [F,a_0] \cdots [F,a_n]) \end{aligned}} \qquad n \text{ even}$$

This is defined when F is at least n-summable and is a cyclic n-cocycle.

2) My approach: I let the gauge group $\mathcal{G} = \mathrm{Map}(M, U_r)$ act on $H^r = (H^+)^r \oplus (H^-)^r$ and consider the map

$$g \longmapsto g F^{\oplus r} g^{-1} = \begin{pmatrix} 0 & g_+ (P^{-1})^{\oplus r} g_-^{-1} \\ g_- P^{\oplus r} g_+^{-1} & 0 \end{pmatrix}.$$

Then the natural invariant forms on the space of F's will pull back to left-invariant forms on $\mathcal{G}$. On the space of invertible operators the basic invariant forms as

$$\operatorname{tr}\left(Q^{-1}dQ\right)^{\text{odd}} . \qquad \left(\text{More precisely } \frac{(-1)^{k-1}(k-1)!}{(2k-1)!} \operatorname{tr}\left(Q^{-1}dQ\right)^{2k-1}\right)$$

Here $Q : H^{+} \longrightarrow H^{-}$ and invariance means with respect to left and right multiplication. We are considering the map $g \longmapsto Q_g = gPg^{-1}$ (abbreviate $P^{\otimes 1}$ to P); then

$$\begin{aligned} Q^{-1}dQ &= \left(gPg^{-1}\right)^{-1}\left(dg\,Pg^{-1} - gPg^{-1}dg\,g^{-1}\right) \\ &= g\,P^{-1}\left[g^{-1}dg, P\right]g^{-1} \end{aligned}$$

i.e. at the identity of $\mathcal{G}$ we have that

$$\iota_X\left(Q^{-1}dQ\right) = P^{-1}[X,P]$$

so we get the left-invariant forms

$$(\star) \qquad \operatorname{tr}\left(P^{-1}[\theta, P]\right)^{\text{odd}} .$$

I guess I need $\qquad F = \begin{pmatrix} 0 & P^{-1} \\ P & 0 \end{pmatrix}$

$$\begin{aligned} \operatorname{tr}\left(\varepsilon F (dF)^{2k-1}\right) &= \operatorname{tr}\left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} & P^{-1} \\ P & \end{pmatrix}\begin{pmatrix} & dP^{-1} \\ dP & \end{pmatrix}\begin{pmatrix} dP^{-1}dP & \\ & dP\,dP^{-1} \end{pmatrix}^{k-1}\right) \\ &= 2 \operatorname{tr} P^{-1}dP\left(dP^{-1}dP\right)^{k-1} \\ &= (-1)^{k-1}\, 2 \operatorname{tr}\left(P^{-1}dP\right)^{2k-1} \end{aligned}$$

So what happens is that when we convert the left-invariant forms $(\star)$ to cyclic cocycles we end up with the same cyclic cocycles as Connes.

Also want

$$\begin{aligned}\varphi_n(a_0,\dots,a_n) &= \operatorname{tr}(\varepsilon\, a_0 [F,a_1]\cdots[F,a_{2k}]) \qquad n=2k\\ &= \tfrac{1}{2}\operatorname{tr}(\varepsilon F [F,a_0]\cdots[F,a_{2k}])\\ &= (-1)^k \operatorname{tr}\left(P^{-1}[P,a_0]\,P^{-1}[P,a_1]\cdots P^{-1}[P,a_{2k}]\right)\end{aligned}$$

<u>Formulas:</u>

$$\begin{aligned}\varphi(a_0,a_1,\dots,a_{2k}) &= \operatorname{tr}(\varepsilon\, a_0 [F,a_1]\cdots[F,a_{2k}])\\ &= \tfrac{1}{2}\operatorname{tr}(\varepsilon F[F,a_0][F,a_1]\cdots[F,a_{2k}])\\ &= (-1)^k \operatorname{tr}\left(P^{-1}[P,a_0]\cdot P^{-1}[P,a_1]\cdots P^{-1}[P,a_{2k}]\right)\end{aligned}$$

$$\tfrac{1}{2}\operatorname{tr}(\varepsilon F (dF)^{2k+1}) = (-1)^k \operatorname{tr}(P^{-1}dP)^{2k+1}$$

$$\begin{aligned}\operatorname{Ind}(ePe) &= (-1)^k\, \tfrac{1}{2}\operatorname{tr}(\varepsilon F[F,e]^{2k+1})\\ &= (-1)^k \operatorname{tr}(\varepsilon\, e[F,e]^{2k})\\ &= \operatorname{tr}\left(P^{-1}[P,e]\right)^{2k+1}\end{aligned}$$

The problem to be solved is to relate these two ways of getting the cyclic cocycles. This apparently requires some understanding of the periodicity process.

I want to describe how to pair an idempotent ~~[illegible]~~ matrix e over A with a left-invariant differential form ω on $\mathcal{G} = \mathrm{Map}(M, U)$. This would explain Connes' approach in terms of mine.

The idempotent defines an element of $K^0(M)$, the form ω defines a coh. class on G. Suppose $\deg(\omega) = 2k+1$. Consider the diagram

$$\begin{tikzcd}
K^0(M) \arrow[r, "\sim"', "\text{periodicity}"] \arrow[d, dotted] & K^0(S^{2k+2} \wedge M) \arrow[r, phantom, "="] & \pi_{2k+2}(BU^M) \arrow[d, equal] \\
HC_{2k}(A) & H_{2k+1}(G) \arrow[l] & \pi_{2k+1}(U^M) \arrow[l, "\text{Hurewicz}"]
\end{tikzcd}$$

where presumably the dotted arrow is Connes' Chern character map. This sends the idempotent e into a constant multiple of $\underbrace{(e\otimes \cdots \otimes e)}_{2k+1} \in C_{2k}(A)$. The constant is rigged so as to be compatible with the S-operator. It would seem to be

$$\frac{1}{2^k \quad 1\cdot 3 \cdots (2k-1)} \quad e^{\otimes(2k+1)}$$

$$= \frac{1}{2} \frac{(k-1)!}{(2k-1)!} \; e^{\otimes(2k+1)}$$

based on chasing through the double complex.

So one problem is to show commutativity of the above diagram. But we haven't made precise the map $H_{2k+1}(G) \longrightarrow HC_{2k}(A)$, which is dual to the process taking a cyclic cocycle to left invariant diffl. form.

Let's review the problem. One start with what Connes calls a Fredholm module i.e. a Hilbert space H which is an A-module and on which one has an involution F such that $[F,a]$ is compact for all $a \in A$. There are graded and ungraded cases.

To fix the ideas suppose M even-diml and $F = \begin{pmatrix} 0 & P^{-1} \\ P & 0 \end{pmatrix}$ where $P : H^+ \longrightarrow H^-$ is a ΨDO of order 0 between two vector bundles.

A Fredholm module which is q-summable determines a sequence cyclic cocycles which I understand best using the LQT theorem as follows. We replace H, F by $H^{\oplus r}$, $F^{\oplus r}$ and let $G = U(r, A)$ act on the new H. Then $g \longmapsto gFg^{-1}$ gives a map from G to the restricted Grassmannian. On this Grassmannian are the standard character forms which pull-back to give left invariant forms on G.

This map $G \longrightarrow$ Grass represents a K-class on G. This is the index of the family of operators $g \longmapsto gFg^{-1}$ parametrized by G in some tautological sense. So the left-invariant forms on G are supposed to represent the character of the index of the family.

(I don't know if there is some problem with the character forms on the restricted Grassmannian representing the character. Ultimately we should take a simple attitude and replace G by a map of a finite-dimensional manifold to G.)

Now the left-inv. diffl. forms on G we have constructed represent primitive cohomology of G, and hence can be detected by homotopy of G, i.e. maps $S \longrightarrow G$ where S is a sphere. By periodicity such maps are given by either idempotent matrices or invertible matrices over A.

So the basic problem for me is to carry out the integration of these differential forms.

To be specific suppose we are in the even case so that M is even, F is graded, and we start with an element α of $K^0(M)$. By

$$K^0(M) = K^0(S^{2k} \wedge M) = [S^{2k} \wedge M, Z \times BU]$$

$$= [S^{2k-1} \wedge M, U] = [S^{2k-1}, G]$$

and my problem is to take the $(2k-1)$-degree form on G and integrate over this S^{2k-1}.

Now let's assume this $(2k-1)$-degree form on G represents the character form of the index.

Consider the diagram

$$\begin{array}{ccc} K^0(M) & \xrightarrow{\sim} & K^0(S^{2k} \wedge M) \\ {\scriptstyle c}\downarrow & & \downarrow{\scriptstyle c} \\ K^0(pt) & \xrightarrow{\sim} & K^0(S^{2k}) \end{array}$$

where c represents capping with the K-homology class represented by the operator F. In this diagram the image of α under $\longrightarrow \downarrow c$

should be the index of the family of operators on the fibres of $S^{2k} \times M$ over S^{2k} which is obtained by mixing F with the bundle over $S^{2k} \times M$ obtained from α.

So it should follow from the commutativity of this diagram and the standard behavior of the Chern character with respect to periodicity that integrating the $(2k-1)$-differential form over the $S^{2k-1} \to G$ belonging to α gives the index of α mixed with F.

Now it should be possible to prove this analytically.

March 17, 1985

It is necessary to get a hold on these F's used by Kasparov and Connes.

I have treated periodicity using finite-dimensional F's. Recall that given a C_k-module V we can consider the space $I_k(V)$ of involutions F on V anti-commuting with $\gamma^1, \dots, \gamma^k$. This is alternately a Grassmannian or unitary grp. $I_0(V)$ is the space of subspaces of V, $I_1(V)$ is the space of unitary maps $V^+ \simeq V^-$.

I now want to relate this periodicity maps using these I_k spaces to the periodicity maps one gets using Clifford modules.

Consider the Bott class in $K_1(\mathbb{C})$. It is represented by the Hopf line bundle on $S^2 = P_1(\mathbb{C})$ and hence by an idempotent 2×2 matrix fn. over $\mathbb{C}$:

$$e = \begin{pmatrix} 1 \\ z \end{pmatrix} \frac{1}{1+|z|^2} \begin{pmatrix} 1 & z^* \end{pmatrix}$$

The image of this is the subbundle $\mathcal{O}(-1)$.

It is interesting to represent the Bott class as a map to Fredholm operators. Over the disk $|z| \leq 1$ one has the Fredholm operator $z \; e_1 e_1^* + e_2 e_2^* + \cdots$ on ℓ^2. This has index zero and is invertible for $z \neq 0$. I want to trivialize this map on $|z| = 1$ using Kuyper's thm. In the present case this is easy.

Let's recall the basic infinite repetition argument

$$\underbrace{a + \overbrace{(-a) + a}} + \underbrace{\overbrace{(-a) + a}}_{} \cdots \quad \begin{array}{l} = a + 0 + \cdots = a \\ = 0 + 0 + \cdots = 0. \end{array}$$

Thus we look at the map from the circle $|z| = 1$ to the unitary group of ℓ^2 given by

$$\begin{pmatrix} z & & & & \\ & z^{-1} & & & \\ & & z & & \\ & & & z^{-1} & \\ & & & & \ddots \end{pmatrix}$$

and join the blocks $\begin{pmatrix} z & \\ & z^{-1} \end{pmatrix}$ to the identity to get the required trivialization.

I recall that the proof of Kuyper's thm. proceeds by showing that [redacted] a map $X \to \tilde{U}$ can be deformed into the subgroup fixing a subspace of infinite dimension and codimension. Then one [redacted] shows the contractibility of the map by the above infinite repetition argument.

[redacted]

A third way to represent the Bott class is via Clifford modules. [redacted] We take the trivial bundle over $\mathbb{C}$ with fibre S_2 and give it the odd degree self-adjoint endomorphism

$$L = x\gamma^1 + y\gamma^2 = \begin{pmatrix} 0 & \bar{z} \\ z & 0 \end{pmatrix}$$

Let's go over the problem under consideration. I start with a Fredholm module (H, F) assoc. say to an operator on M. To fix the ideas suppose we are in the even case. Then attached to (H, F) are cyclic cocycles φ_{2k} k sufficiently large. I define them as left-invariant forms on $\mathcal{G} = Map(M, U(n))$ and apply the LQT-theory to write them in terms of cyclic cocycles.

What I learned today:

Let's adopt Connes' viewpoint relative to which a Fredholm ~~operator~~ module F is to be regarded as a reduction of an invertible Fredholm ~~operator~~ module. Then one can define the family of cyclic cocycles attached to a Fredholm modules. They are left-invariant differential forms on $\mathcal{G} = \mathrm{Map}(M, U)$.

Let's take then the even case and the cyclic cocycle φ_{2k}. This corresponds to a form of degree $2k+1$ on $\mathcal{G}$.

Next I take e representing an element of $K_0(M)$ and use periodicity to associate to e a map $S^{2k+1} \times M \longrightarrow U$, i.e. a map $S^{2k+1} \longrightarrow \mathcal{G}$. Then I can integrate the diff'l form over this sphere.

To see what we get we use the fact that we might as well reduce by the idempotent e first. Thus one can suppose $e = 1$ by replacing the Fredholm module by its reduction relative to e. (Define this by taking a direct sum of copies, but then use the A module structure on the image of e and 0 on the image of $1-e$.)

But if $e = 1$, then the map $S^{2k+1} \times M \longrightarrow U$ factors $S^{2k+1} \times M \xrightarrow{pr_1} S^{2k+1} \longrightarrow U$, where the second map is the Bott map. ($S^{2k+1} \longrightarrow U(2^k)$?). Thus $S^{2k+1} \longrightarrow \mathcal{G}$ factors $S^{2k+1} \longrightarrow U \longrightarrow \mathcal{G}$ where the second is the inclusion of the constant gauge transformations.

So one takes the cyclic cocycle over A and restricts in to $\mathbb{C}$ where it becomes a multiple of the unique cyclic cocycle of that degree. It follows that if things are normalized correctly the integral of the differential form will be just the index.

March 18, 1985

Let's go over the AS proof of periodicity based on Kuiper's theorems. It would be nice to understand why it really works. But I should be able to link it with the Bott periodicity maps constructed above in finite dimensions.

Outline of my previous understanding of this proof.

$\mathcal{a} = \mathcal{B}/\mathcal{K}$ is the Calkin algebra

$\mathcal{F}$ = Fredholm operators = inverse image of $\mathcal{a}^*$ in $\mathcal{B}$. We know $\mathcal{F}$ deforms to the subgroup of essentially unitary operators, i.e. inverse image of $\mathcal{U}(\mathcal{a})$.

Similarly $\mathcal{F}_1$ = self-adjoint Fredholm operators w essential spectrum on both sides of 0 deforms to the self-adjoint operators having essential spectrum ± 1. This maps onto $\mathcal{I}(\mathcal{a})$ = space of non-trivial idempotents in $\mathcal{a}$.

The first step is the exact sequences

$$1 \longrightarrow \mathcal{U} \longrightarrow \tilde{\mathcal{U}} \longrightarrow \mathcal{U}(\mathcal{a})_0 \longrightarrow 1$$

$$1 \longrightarrow \mathcal{U}(\mathcal{a})_0 \longrightarrow \mathcal{U}(\mathcal{a}) \xrightarrow{\text{index}} \mathbb{Z} \longrightarrow 1$$

and because of Kuiper's thm. that $\tilde{\mathcal{U}} \sim pt$, one get

$$B\mathcal{U} \sim \mathcal{U}(\mathcal{a})_0$$

$$\mathbb{Z} \times B\mathcal{U} \sim \mathcal{U}(\mathcal{a}) \sim \mathcal{F}$$

The next step is

$$\mathcal{I}(\mathcal{a}) = \mathcal{U}(\mathcal{a})^* / \mathcal{U}(\mathcal{a})^* \times \mathcal{U}(\mathcal{a})^* \sim B\,\mathcal{U}(\mathcal{a})$$

which should be a [redacted] version of the standard proof that BU = Grassmannian. I don't know if Kuiper's result is used here.

The key step is [redacted] to show that the map

$$\mathcal{F}_1 \longrightarrow \mathcal{U} \quad , \quad A \longrightarrow -\exp(i\pi A)$$

is a homotopy equivalence. This is [redacted] uses the natural filtration by number of zero (resp. -1) eigenvalues + Kuiper's thm. to identify the homotopy types of the strata. The idea is that the ± 1 spaces for an operator in $\mathcal{F}_1$ become the $+1$ eigenvalues of an operator in $\mathcal{U}$. Thus we are concerned with the space of splittings of Hilbert space, i.e. with

$$\tilde{\mathcal{U}} / \tilde{\mathcal{U}} \times \tilde{\mathcal{U}}$$

which is contractible by Kuiper's thm. *

Let's now try to fit this [redacted] together with our earlier discussion about periodicity maps using Clifford algebras.

---

* Idea: We know that by looking at the eigenvalues of a [redacted] self-adjoint operator in finite dimensions that we get the simplicial structure of the building. There's a possibility that the simplicial structure is involved in cyclic cohomology.

Recall that if V is an ungraded C_k-module with inner product we put

$$I_k(V) = \{ J \in \mathrm{End}(V) \mid J^2 = 1,\ J = J^*,\ \gamma^a J + J\gamma^a = 0 \}$$

so that

$$I_0(V) = \mathrm{Grass}(V)$$
$$I_1(V) = \mathrm{Unitary\,Isoms}(V^+, V^-)$$

We want to consider the case where V is a Hilbert space and J is Fredholm.

So I am being led to look at the K-theory of C_k in the Kasparov sense:

We know graded C_k-modules which are fin. diml do not give $K(\mathbb{R}^k)$, but the do in the Kasparov game. How?

$k=0$. A graded C_0-module is a graded v.s. $V = V^+ \oplus V^-$, and the naive K-theory of these is $\mathbb{Z}+\mathbb{Z}$. In the KK game one looks at $H = H^+ \oplus H^-$ (both inf. diml in the stable model) equipped with and $F = \begin{pmatrix} 0 & Q \\ P & 0 \end{pmatrix}$ of odd degree such that $F^2 - 1$ is compact. And one works modulo homotopy, ~~getting~~ $\mathbb{Z}$ for the K-theory.

$k=1$. A graded C_1-module is equivalent to an ungraded v.s. The naive K-group is $\mathbb{Z}$. In the KK game one looks at F on a Hilbert space H such that $F^2 - I$ is compact. [redacted] Working stably both the $F=1$ and $F=-1$ spaces are infinite diml. Working mod homotopy the K-group is zero.

What are the periodicity maps? We want to use the same maps as before. I think we want to exploit the idea that there are three kinds of F's - Fredholm, unitary, unitaries $\equiv I \mod \mathcal{K}$.

Let us now consider the periodicity map

$$\mathcal{J}(a) \longrightarrow \Omega\,\mathcal{U}(a)$$

which recall takes an involution J into the paths from 1 to -1 given by

$$(\cos\theta) + i(\sin\theta)J \qquad 0 \leq \theta \leq \pi.$$

Where does this come from? We are really dealing with the map from $k=1$ to loops on $k=0$. A graded C_1-module with F has the form $H = H^+ \oplus H^-$ where $H^+ = H^-$ and

$$\varepsilon = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \qquad \gamma^1 = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} \qquad F = \begin{pmatrix} & -iJ \\ iJ & \end{pmatrix} = \gamma^2 J$$

and where $J^2 \equiv 1 \mod \mathcal{K}$. Then

$$\cos\theta\,\gamma^2 + \sin\theta\,F = \begin{pmatrix} 0 & \cos\theta - i\sin\theta\,J \\ \cos\theta + i\sin\theta\,J & 0 \end{pmatrix}$$

so we get the essentially unitary operator

$$\cos\theta + i(\sin\theta)J.$$

Note that if $J^2 = 1$ then

$$e^{i\theta J} = \cos(\theta J) + i\sin(\theta J)$$
$$= \cos\theta + i\sin(\theta)J$$

Thus we have the periodicity map

$$\mathcal{I}(a) \longrightarrow \Omega(\mathcal{U}(a);1,-1)$$

which we want to show is a homotopy equivalence. On the other hand the fibration

$$\mathcal{U} \longrightarrow \tilde{\mathcal{U}} \longrightarrow \mathcal{U}(a)_0$$

with the contractibility of $\tilde{\mathcal{U}}$ determines a map

$$\Omega(\mathcal{U}(a)_0,1,-1) \longrightarrow -\mathcal{U}$$

by lifting paths, which is a heq.

Now over [redacted] $\mathcal{F}_1$ we can do this path lifting explicitly. Given $A \in \mathcal{F}_1$ we can consider the path $e^{i\theta A}$, $0 \le \theta \le \pi$, in $\tilde{\mathcal{U}}$. This starts at 1 and covers $e^{i\theta J} = \cos\theta + i(\sin\theta)J$ in $\mathcal{U}(a)_0$ where $A \mapsto J$. Thus the endpoint map is $A \longmapsto e^{i\theta A}$ which is [redacted] in $-\mathcal{U}$. So

$$\begin{tikzcd}
\mathcal{F}_1 \arrow[r, "\exp"] \arrow[d, "\sim"'] & -\mathcal{U} \\
\mathcal{I}(a) \arrow[r] & \Omega(\mathcal{U}(a)_0;1,-1) \arrow[u, "\sim"']
\end{tikzcd}$$

commutes up to homotopy. We conclude then that the Bott periodicity map is an heq iff exp is.

So far we have the part $\mathcal{F}_1 = \Omega\mathcal{F}_0$ of periodicity. We next look at $k=2$.

March 20, 1985

$$\mathcal{F}_0 = \{ F = \begin{pmatrix} 0 & P^* \\ P & 0 \end{pmatrix} \mid F^2 - 1 \in \mathcal{K} \}$$

$$= \{ P \in \mathcal{B}(H) \mid PP^* \equiv P^*P \equiv 1 \mod \mathcal{K} \}$$

$$\mathcal{F}_1 = \{ P \in \mathcal{B}(H) \mid \begin{matrix} P^2 \equiv 1 \\ P = P^* \end{matrix} \mod \mathcal{K} \}$$

are the two spaces taking part in periodicity. They map onto the unitary elements (resp. involutions [redacted]) in the Calkin algebra $\mathcal{A}$ with contractible fibres.

[redacted] I want now to set up the periodicity maps. One idea is to consider the fibrings

$$U \longrightarrow U(\mathcal{B}) \longrightarrow U(\mathcal{A})$$

$$Gr \longrightarrow \mathcal{I}(\mathcal{B}) \longrightarrow \mathcal{I}(\mathcal{A})$$

defined as follows. In the second case we consider self adjoint [redacted] involutions in $\mathcal{B}$ and $\mathcal{A}$ and the fibre Gr of involutions J in $\mathcal{B}$ which are congruent mod $\mathcal{K}$ to a fixed involution. This is the restricted Grassmannian. In the first case we do the same thing but using odd involution relative to a grading. Thus we take $F = \begin{pmatrix} 0 & g^{-1} \\ g & 0 \end{pmatrix}$ over $\mathcal{A}$ and try to lift to an F in $\mathcal{B}$ which is possible over the index zero component. The fibre is then the unitaries congruent to 1 mod $\mathcal{K}$.

The next step is to consider the Bott maps down in $\mathcal{A}$. In the second case we have the map

$$\mathcal{U}(\mathcal{A}) \longrightarrow \Omega(\mathcal{I}(\mathcal{A});\ \varepsilon, -\varepsilon)$$

$$F = \begin{pmatrix} 0 & g^{-1} \\ g & 0 \end{pmatrix} \longmapsto F_\theta = (\cos\theta)\varepsilon + (\sin\theta)F$$
$$= \begin{pmatrix} \cos\theta & (\sin\theta)g^{-1} \\ (\sin\theta)g & -\cos\theta \end{pmatrix}$$

which we know associates to g the graph of $(\tan\frac{\theta}{2})g$ as θ runs over $[0,\pi]$.

We want to lift a loop in $\mathcal{I}(\mathcal{A})$ into $\mathcal{I}(\mathcal{B})$ and then take the endpoint getting a map

$$\mathcal{U}(\mathcal{A}) \longrightarrow Gr$$

defined up to homotopy. We can do this explicitly by replacing $\mathcal{U}(\mathcal{A})$ by $\mathcal{F}_0$ and using

$$F = \begin{pmatrix} 0 & p^* \\ p & 0 \end{pmatrix} \longmapsto F_\theta = e^{\frac{\theta}{2}F\varepsilon}\,\varepsilon\, e^{-\frac{\theta}{2}F\varepsilon} = e^{\theta F\varepsilon}\cdot\varepsilon.$$

What one is doing is to regard $\mathcal{I}(\mathcal{B})$ as $\mathcal{U}(\mathcal{B})/\text{centralizer of } \varepsilon$ and using the exponential map in $\mathcal{U}(\mathcal{B})$ restricted to the subspace of the Lie algebra complementary to the centralizer of ε

Let me try to understand things a bit better. The thing about the Atiyah-Singer proof which I missed before is the fact that the restricted Grassmannian has the same kind of eigenvalue description as the restricted unitary group. This should have been obvious if we think of the Grassmannian as a symmetric space.

We can look at things as follows. Let ε be our given (s.a.) involution on V and let the unitary group U act on these involutions. Thus we have

$$U/U^+ \times U^- \hookrightarrow \mathcal{I}(V)$$
$$g \longmapsto g\varepsilon g^{-1}.$$

We split the Lie algebra $\mathfrak{g}$ of U into $k + p$ as usual and then consider the exponential map

$$\exp : p \longrightarrow \text{Grass}.$$

This should be onto (at least in f.dims.). Now if $g = e^X$, $X \in p$ then $\varepsilon g \varepsilon = g^{-1}$ so $g\varepsilon g^{-1} = g^2 \varepsilon$.

Notice that for any $g \in U$

$$(g\varepsilon)^2 = g\varepsilon g\varepsilon = 1 \iff \varepsilon g \varepsilon = g^{-1}.$$

Thus we can identify the Grassmannian with the subset of U consisting of g with $\varepsilon g \varepsilon = g^{-1}$.

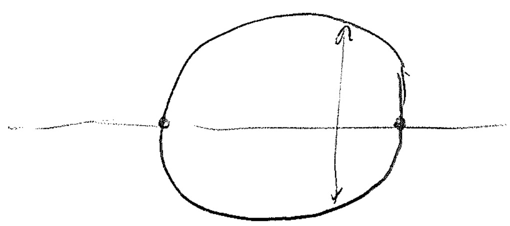
~~Proof: The space of involutions ... are conjugate ... fix of involution~~

Proof: Consider the set of involutions J on V. Then $J \longmapsto J\varepsilon$, $g \longmapsto g\varepsilon$

is a bijection of this space with the subset of $\mathcal{U}$ consisting of g with $\varepsilon g \varepsilon = g^{-1}$.

<u>Proof:</u> If $\varepsilon g \varepsilon = g^{-1}$, then $(g\varepsilon)^2 = 1$ so $g\varepsilon$ is an involution, and its unitary as g, ε are. If J is a unitary involution, then $(J\varepsilon)^{-1} = \varepsilon J = \varepsilon (J\varepsilon)\varepsilon$.

So to describe the Grassmannian one takes unitaries g such that $\varepsilon g \varepsilon = g^{-1}$. This implies that the eigenvalues are stable under conjugation.

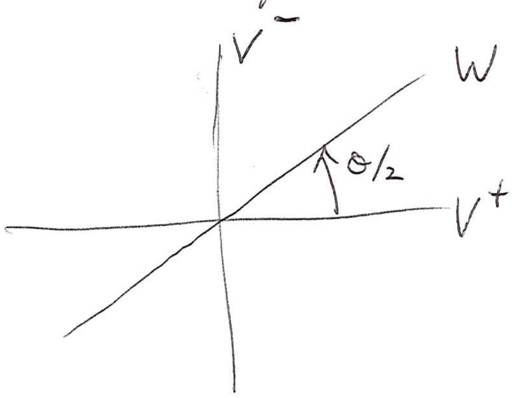


So you see the spectral structure rather nicely. The eigenvalues not in ± 1 will have eigenspaces paired via ε. The $+1$ eigenspace will be stable under ε; also the -1 eigenspace.

I now want to work out the geometric meaning of these eigenvalues. Thus I start with a fixed involution ε on V. Then by the above, the different Grassmannians of V appear as components of $\{g \in \mathcal{U} \mid \varepsilon g \varepsilon = g^{-1}\}$. For example if $\varepsilon = 1$, then we get $\{g \in U \mid g^2 = 1\}$. But this isn't interesting because we really want to consider the case where ε belongs to the Grassmannian.

In any case we can look at the spectral structure to get some understanding. ε defines a grading $V = V^+ \oplus V^-$. Clearly given $W \subset V$ we want to project it onto V^+ and V^- to get decompositions of these; this must correspond to the ± 1 eigenvalues of g. Leave these for the moment and consider the case where W sets up an isomorphism between V^+ and V^-.

Let's first take $V^+ = V^- = \mathbb{C}$.



The inv. corresp. to W is
$$\begin{pmatrix} \cos\theta & +\sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} = J_\theta$$

so the corresponding unitary is
$$J_\theta \varepsilon = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

This is still not as clear as I would like, because I don't see the unitary transformation very well. The actual eigenspaces are in some complex direction.

March 21, 1985

Witten's way to attach a subspace to an $F = \begin{pmatrix} 0 & T^* \\ T & 0 \end{pmatrix}$ is to consider the "massive Dirac op"

$$D = m\varepsilon + F = \begin{pmatrix} m & T^* \\ T & -m \end{pmatrix} \qquad \text{which satisfies } D^2 = m^2 + F^2 \geq m^2$$

and take the decomposition into the positive and negative eigenspaces. To find this we compute the involution:

$$J = D/|D| = \begin{pmatrix} m & T^* \\ T & -m \end{pmatrix} \begin{pmatrix} m^2 + T^*T & 0 \\ 0 & m^2 + TT^* \end{pmatrix}^{-1/2}$$

$$= \begin{pmatrix} m(m^2+T^*T)^{-1/2} & T^*(m^2+TT^*)^{-1/2} \\ T(m^2+T^*T)^{-1/2} & -m(m^2+TT^*)^{-1/2} \end{pmatrix}$$

then form the projector

$$e = \frac{J+1}{2} = \begin{pmatrix} \dfrac{m(m^2+T^*T)^{-1/2}+1}{2} & \dfrac{T^*(m^2+TT^*)^{-1/2}}{2} \\ \dfrac{T(m^2+T^*T)^{-1/2}}{2} & \dfrac{-m(m^2+TT^*)^{-1/2}+1}{2} \end{pmatrix}$$

which should have the positive eigenspace for its image, and then take the image of H^+. We get (at least if $m \neq 0$)

$$\operatorname{Im} \begin{pmatrix} m(m^2+T^*T)^{-1/2}+1 \\ T(m^2+T^*T)^{-1/2} \end{pmatrix} = \operatorname{Im} \begin{pmatrix} m + (m^2+T^*T)^{1/2} \\ T \end{pmatrix}$$

$$= \text{graph of} \quad T\left(m + (m^2+T^*T)^{1/2}\right)^{-1}$$

Suppose T unitary. Then $x = \left(m + (m^2+T^*T)\right)^{-1} = \left(m + \sqrt{m^2+1}\right)^{-1}$ $= -m + \sqrt{m^2+1}$ goes from $+\infty$ to 0 as m goes from $-\infty$ to $+\infty$.

Think in terms of the periodicity map $(F^2=1)$

$$\cos\theta\,\varepsilon + (\sin\theta)F \qquad 0 \leq \theta \leq \pi$$

which goes from ε to $-\varepsilon$. Now I consider the involution

$$\frac{D}{|D|} \quad \text{where} \quad D = m\varepsilon + F$$

and you see that m goes from $+\infty$ to $-\infty$. Finally you can consider the graph version

$$\operatorname{Im}\begin{pmatrix} \frac{1+\cos\theta}{2} \\ \frac{\sin\theta}{2}T \end{pmatrix} = \operatorname{Im}\begin{pmatrix} 1 \\ xT \end{pmatrix}$$

where $x = \tan\frac{\theta}{2}$ goes from 0 to ∞.

In each case I think of a path from ε to $-\varepsilon$ or from H^+ to H^-.

Now we decompose according to the eigenvalues of T which as we have seen are real nos. $\lambda \geq 0$. Then there is a difference as to the actual correspondence between the parameters. Say $T = \lambda > 0$. Then

$$\cancel{(\cos\theta)\varepsilon + (\sin\theta)a = \begin{pmatrix} \cos & a\sin \\ a\sin & -\cos\theta \end{pmatrix}}$$

I have to go over the three possibilities.

1) graph $\begin{pmatrix} [\text{redacted}] \\ x\lambda \end{pmatrix}$ $\qquad 0 \leq x < \infty$

2) Witten $=$ graph $\left(\dfrac{\lambda}{m+\sqrt{m^2+\lambda^2}}\right)$

3) $e^{[\text{redacted}]\theta F\,\varepsilon}\,[\text{redacted}]\,\varepsilon = \begin{pmatrix} \cos(\theta\lambda) & +\sin(\theta\lambda) \\ \sin(\theta\lambda) & -\cos(\theta\lambda) \end{pmatrix}$ gives graph $\left(\tan\frac{\theta\lambda}{2}\right)$

So the speed of the correspondence depends upon λ.

Now what I have to do is to understand whether the periodicity map can be done using these variants. The essential object is the null-spaces of T and T^*.

So let's see what happens to our path when we encounter a zero eigenvalue for $\begin{pmatrix} 0 & T^* \\ T & 0 \end{pmatrix}$. Thus we take $\lambda = 0$.

1) In the graph construction we get simply the constant path at H^+.

2) In the Witten case we look at the positive eigenspace of $\begin{pmatrix} m & 0 \\ 0 & -m \end{pmatrix}$ and this jumps from H^+ to H^- as m passes from >0 to <0.

3) Constant path at H^+.

Is it possible that the graph could give a good periodicity map?

March 23, 1985

Let $f: \mathbb{N}_{>0} \longrightarrow \mathbb{R}$ satisfy

$$f(m+n) \geq f(m) + f(n).$$

Extend f to $\mathbb{N}$ by letting $f(0)=0$; this inequality still holds. Use the division algorithm:

$$n = qd + r \qquad 0 \leq r < d, \qquad q = \left[\tfrac{n}{d}\right] \geq 0$$

Then by induction

$$f(n) \geq \left[\tfrac{n}{d}\right] f(d) + f(r)$$

So

$$\frac{f(n)}{n} \geq \frac{1}{n}\left[\frac{n}{d}\right] f(d) + \frac{f(r)}{n}$$

and letting $n \to \infty$ we get

$$\varliminf_n \frac{f(n)}{n} \geq \underbrace{\left(\lim_n \frac{1}{n}\left[\frac{n}{d}\right]\right)}_{1/d} f(d) + \underbrace{\lim_n \frac{f(r)}{n}}_{0}$$

$$= \frac{f(d)}{d}$$

So $\displaystyle\varliminf_n \frac{f(n)}{n} \geq \varlimsup_d \frac{f(d)}{d}$ and we see that $\displaystyle\lim \frac{f(n)}{n}$ exists (possibly it is $+\infty$).

<u>Lemma</u>: If $f: \mathbb{N}_{>0} \longrightarrow \mathbb{R}$ satisfies $f(m+n) \geq f(m) + f(n)$ then $\displaystyle\lim_{n\to\infty} \frac{f(n)}{n}$ exists in $(-\infty, \infty]$.

<u>Application</u>: Let μ be a prob. measure in a vector space V and μ_n the image of $\mu^{\otimes n}$ on V^n under $(v_1, \ldots, v_n) \longmapsto \frac{1}{n}\sum v_i$. Then given a convex set

$A \subset V$ we will show

$$\lim \frac{1}{n} \log \mu_n(A) \quad \text{exists in } (-\infty, 0]$$

We have

$$\mu_n(A) = \int_{\frac{1}{n}\sum x_i \in A} \mu(dx_1) \cdots \mu(dx_n) = \mu^{\otimes n}(S_n)$$

where $S_n = \{ (x_i) \in V^n \mid \frac{1}{n}\sum x_i \in A \}$. We have

$$S_m \times S_n \subset S_{m+n}$$

since if $\frac{1}{m}\sum_{i=1}^{m} x_i \in A$, $\frac{1}{n}\sum_{j=1}^{n} x_{m+j} \in A$, then

$$\frac{1}{m+n} \sum_{i}^{m+n} x_i = \frac{m}{m+n} \frac{1}{m} \sum_{1}^{m} x_i + \frac{n}{m+n} \frac{1}{n} \sum_{j}^{n} x_{m+j}$$

$$\in \frac{m}{m+n} A + \frac{n}{m+n} A \subset A$$

as A is convex. Thus

$$\mu_{m+n}(A) = \mu^{\otimes(m+n)}(S_{m+n}) \geqslant \mu^{\otimes(m+n)}(S_m \times S_n)$$

$$= \mu^{\otimes m}(S_m)\, \mu^{\otimes n}(S_n) = \mu_m(A)\mu_n(A)$$

and so if

$$f(n) = \log \mu_n(A)$$

f satisfies $f(m+n) \geqslant f(m) + f(n)$ and we can apply the preceding lemma.

Except $f(n)$ might be $-\infty$.

March 24, 1985

Fatou's lemma: If $f_n \geq 0$, then

$$\int \liminf f_n \leq \liminf \int f_n$$

Proof: Recall first the monotone convergence thm: If $f_n \geq 0$, $f_n \leq f_{n+1} \leq \cdots$ then one has

$$\sup \int f_n = \int \sup f_n .$$

This is because

$$\int f_n = \mu \{ (x,y) \in X \times \mathbb{R}_{\geq 0} \mid 0 \leq y < f_n(x) \}$$

$$= \mu \,(\text{subgraph } f_n)$$

and $$\text{subgraph}\,(\sup f_n) = \bigcup_n \text{subgraph } f_n$$

and $$\mu\left(\bigcup A_n \right) = \sup \mu(A_n) \qquad \text{for } A_n \subset A_{n+1} \subset \cdots$$

(Notice this is not true for decreasing sequences unless $\mu(A_n) < \infty$.)

Then

$$\underbrace{\int \inf_{n \geq k} f_n}_{\text{increasing in } n} \leq \inf_{n \geq k} \int f_n$$

so applying monotone convergence

$$\int \liminf f_n = \int \sup_k \inf_{n \geq k} f_n = \sup_k \int \inf_{n \geq k} f_n$$

$$\leq \sup_k \inf_{n \geq k} \int f_n = \liminf \int f_n$$

Application

$$Z(J) = \int e^{(J,x)}\, \mu(dx) \quad \in \quad [0,\infty]$$

If $J_n \longrightarrow J$ then we have

$$\int \liminf e^{(J_n,x)} \leq \liminf \int e^{(J_n,x)}$$

or

$$Z(J) \leq \liminf_n Z(J_n)$$

Since log is continuous and monotone it preserves inf's + sup's, so we have

$$\log Z(J) \leq \liminf_n \log Z(J_n)$$

Next recall the definition

$$W(x) = \sup_J\, (J,x) - \log Z(J)$$

is a sup of continuous functions so it has a semi-continuity property. To express this we look at the supergraph $f = \{(x,y) \mid y \geq f(x)\}$. We ask what it means for this supergraph to be closed. If it is closed, then for $x_n \longrightarrow x$ one has that $(x, \liminf f(x_n))$ is a limit point of the supergraph, so

$$\liminf f(x_n) \geq f(x)$$

called <u>lower semi-continuous</u>

Conversely if this holds and $x_n \to x$, $y_n \to y$ with $y_n \geq f(x_n)$, then

$$y = \lim y_n \geq \liminf f(x_n) \geq f(x)$$

so (x,y) belongs to the supergraph of f.

Summary: f is lower semi-continuous means

$$\text{a)}\quad \liminf f(x_n) \geq f(x) \qquad \text{if } x_n \to x$$
$$\Updownarrow$$
$$\text{b)}\quad \{(x,y) \mid y \geq f(x)\} \text{ is closed.}$$

From Fatou's lemma + [redacted] inequality

$$\left|\int fg\right| \leq \|f\|_p \|g\|_q \qquad \frac{1}{p} + \frac{1}{q} = 1$$

which comes from convexity of the exponential function we know that

$$\log Z(J) \qquad\qquad Z(J) = \int e^{(J,x)}\,\mu(dx)$$

has closed convex supergraph. Hence this fn. is convex and l.s.c.

This implies that $\{J \mid Z(J) < \infty\}$ is convex. In the following we assume it is non-empty (i.e. usually that $Z(0) = 1$). Then when we take the Fenchel transform

$$W(x) = \sup_J \, (J,x) - \log Z(J)$$

we have $W(x) \geq (J_0, x) - \log Z(J_0)$. Otherwise $W(x) = -\infty$ for all x.

I want now to show that I can recover $\log Z(J)$ from $W(x)$:

$$\log Z(J) = \sup_x \, (J,x) - W(x)$$

Here's the motivation from yesterday:

Let's start with a probability measure μ on $V = \mathbb{R}^n$ say with compact support. We assume it is not supported in a hyperplane. Then we know that the map

$$J \longmapsto x_J = \nabla \log Z(J) \quad , \quad V' \longrightarrow V$$

is a local diffeomorphism. (We know that the Jacobian matrix at $J \in V'$, which is a linear map $V' \longrightarrow V$ or equiv. a quadratic form on V', is positive definite. It gives the variance of $e^{Jx}\mu / Z(J)$.)

I want to prove that this map is a diffeomorphism of V' with the interior of the convex hull of the support of μ. This result is like the Atiyah–Guillemin–Sternberg convexity thm.

I can define the Fenchel transform $W(x)$ as above. The first thing is to check that it coincides with the Legendre transform at a point of the form x_J. So let $x = x_{J_0}$ be the gradient of $\log Z$ at J_0. The smooth function

$$J \longmapsto (J, x) - \log Z(J)$$

has negative definite second derivative everywhere and zero first derivative at $J = J_0$. So it follows from calculus that this function has a maximum at $J = J_0$. Thus we see that

$$W(x) = (J_0, x) - \log Z(J_0)$$

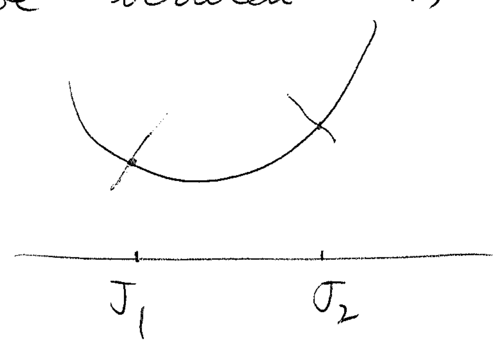
if $x = x_{J_0}$.

But now recall that where the Legendre transform is defined one has

$$DW(x) = J \qquad \text{when} \qquad x = x_J .$$

So this shows that $J \mapsto x_J$ is 1-1, and so this map is a diffeomorphism with its image which is an open subset of V. I would like to identify this open subset with the interior of the set where W is finite.

Actually the fact that $J \mapsto x_J$ is an embedding is much simpler to see. We know $\log Z(J)$ is strictly convex so given two points J_1, J_2 restrict this function to the line containing J_1, J_2. Then the gradient of $\log Z(J)$ has to move between J_1, J_2



otherwise over this line $\log Z(J)$ would be linear.

Fenchel transform stuff. We are interested in subsets of $V \times \mathbb{R}$ of the form

$$\hat{Y} = \{ (x,a) \mid a \geq J \cdot x - c \,,\ \forall (J,c) \in Y \subset V' \times \mathbb{R} \}$$

Let's fit this into a Galois correspondence setup. [redacted] We consider the relation $Z \subset (V \times \mathbb{R}) \times (V' \times \mathbb{R})$

$$Z = \{ (x, a, J, c) \mid \quad a + c \geq J \cdot x \}$$

Given $Y \subset V' \times \mathbb{R}$, then $\hat{Y} \subset V \times \mathbb{R}$ is the subset of all (x, a) related by Z to every elt of Y. We have

$$Y \subset Y_1 \implies \hat{Y} \supset \hat{Y}_1 .$$

Also
$$\hat{X} \supset Y \qquad \text{(means any } \eta \in Y \text{ is } Z\text{-related to any } \xi \in X)$$
$$\Updownarrow$$
$$X \subset \hat{Y} \qquad \text{(means any } \xi \in X \text{ is related to any } \eta \in Y)$$

Thus $\quad \hat{X} \subset \hat{X} \implies \hat{\hat{X}} \supset X$; and $\hat{\hat{\hat{X}}} \subset \hat{X}$. Thus $\hat{X} = \hat{\hat{\hat{X}}}$ and we have a 1-1 [redacted] order reversing correspondence between closed subsets of $V \times \mathbb{R}$ and $V' \times \mathbb{R}$.

Next we have to identify the closed subsets of $V \times \mathbb{R}$ and $V' \times \mathbb{R}$ for this closure relation.

$$\hat{Y} = \left\{ (x, a) \;\middle|\; \text{for all } (J, c) \in Y \text{ we have } \begin{array}{l} a + c \geq J \cdot x \\ Jx - a \leq c \end{array} \right\}$$

is a closed convex subset of $V \times \mathbb{R}$ such that $(x, a) \in \hat{Y}$ $\implies (x, a') \in \hat{Y}$ for $a' \geq a$.

At this point I need to know that a convex closed set is the intersection of half-spaces. This is proved by some Hahn-Banach type argument, i.e. by induction on the dimension in the case of finite dims. Let's review this.

March 25, 1985

Fenchel transform.

Let V be a real vector space of finite dimension, let V' be its dual. Set

$$X = V \times \mathbb{R} \qquad Y = V' \times \mathbb{R}$$

and let $\Gamma \subset X \times Y$ be

$$\Gamma = \{ (v, a, \xi, c) \mid a + c \geq \xi . v \}$$

Given $A \subset X$ define

$$A' = \{ y \in Y \mid A \times \{y\} \subset \Gamma \}$$

$$= \text{largest } B \subset Y \text{ such that } A \times B \subset \Gamma.$$

Define $B' \subset X$ for $B \subset Y$ similarly. Then clearly

$$A_1 \subset A_2 \implies A_1' \supset A_2' \ ; \quad \text{similarly for } B$$

$$A \subset B' \iff A \times B \subset \Gamma \iff B \subset A'.$$

Hence $A' \subset A' \implies A \subset A''$; similarly $B \subset B''$.
Thus $A' \subset A'''$ and $A' \supset A'''$ so $A' = A'''$; sim. for B.

Therefore we have a Galois correspondence between subsets of X of the form B' and subsets of Y of the form A'. Also can say subsets closed for the closure operation $A \mapsto A''$; sim. for B.

Now our problem is to identify the "closed" sets in the case at hand. First note that if $B = \phi$, then $B' = X$, whereas if $B \neq \phi$, say B contains (ξ, c), then

$$B' \subset \{(\xi, c)\}' = \underbrace{\{ (v, a) \mid a + c \geq \xi . v \}}_{\text{supergraph of the linear fn. } v \mapsto \xi . v - c}$$

In fact

$$B' = \{ (v,a) \mid \forall (\xi,c) \in B, \; a \geq \xi . v - c \}$$

$$= \{ (v,a) \mid a \geq \sup_{(\xi,c)\in B} \xi . v - c \}$$

is the supergraph of

$$f(v) = \sup_{(\xi,c)\in B} \xi . v - c$$

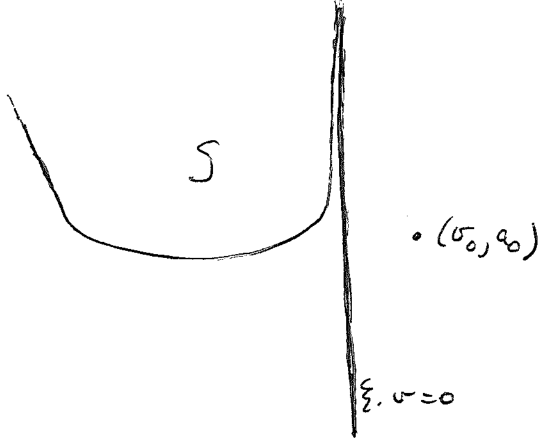
Note that for $B \neq \phi$ this function has values in $\mathbb{R} \cup \{+\infty\}$ (and for $B = \phi$ it has the constant value $-\infty$). Because the supergraph is convex + closed, it follows that the function f is convex and lower semi-continuous.

Conversely given $f : V \longrightarrow \mathbb{R} \cup \{+\infty\}$ convex and l.s.c., its supergraph S is closed and convex. If $(v_0, a_0) \notin S$, we can separate S from this point by a hyperplane

Case 1. The hyperplane is of the form $a = \xi . v - c$.

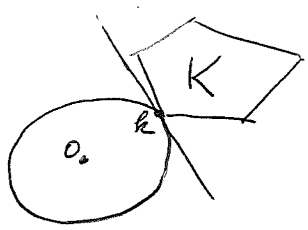
Then $\quad S \subset \{ (v,a) \mid a \geq \xi . v - c \}, \qquad a_0 < \xi . v_0 - c$

Case 2. The hyperplane is of the form $\xi . v = 0$ with $\xi \neq 0$. In this case we have the picture

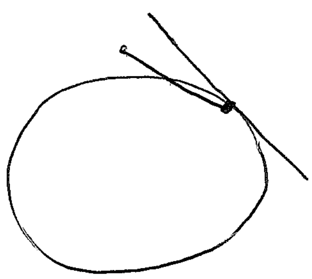


Somehow we are going to have to jiggle this hyperplane.

It seems necessary to understand the separation thm. In finite dimensions we can use an inner product. Given K closed convex a basic result says K has a unique point k closest to O.



It follows that K lies outside the tangent hyperplane to the sphere around O passing thru k. (Otherwise you could join a point inside the hyperplane to k and get a contradiction:



Now let's go back to S and (v_0, a_0) as above. We draw in the largest open ball around a_0 not meeting S, and look at the tangent plane at the unique intersection pt of the boundary with S. If this tangent plane is vertical then we do the same for $a_0 - t$. As t increases we watch the radius, i.e. the distance of $(v_0, a_0 - t)$ from S. The distance has to begin to increase at some point; otherwise S would contain a vertical line. At that point the tangent hyperplane no longer stays vertical:

and by continuity we will get a hyperplane separating (v_0, a_0) from S which isn't vertical.

So therefore we see that the supergraph of an $f: V \longrightarrow \mathbb{R} \cup \{+\infty\}$ which is convex and l.s.c. is the intersection of the upward pointing half spaces containing it. This implies it is closed for the Galois correspondence.

Now we are ready to set up the Fenchel transform. We consider the $A \subset X$ such that $A = A''$ and such that $A \neq \phi, X$. Then A is the supergraph of a unique conv. l.s.c. fn. $f: V \longrightarrow \mathbb{R} \cup \{+\infty\}$ such that $f(x) < \infty$ for some x. Then

$$A = \{ (v, a) \mid a \geqslant f(v) \}$$

$$A' = \{ (\xi, c) \mid \forall (v, a),\ a \geqslant f(v) \Rightarrow a \geqslant \xi . v - c \}$$

$$= \{ (\xi, c) \mid \forall v,\ f(v) \geqslant \xi v - c \}$$

$$= \{ (\xi, c) \mid \forall v, \quad c \geq \xi\cdot v - f(v) \}$$

$$= \{ (\xi, c) \mid c \geq \sup_v \xi v - f(v) \}$$

Thus if A = supergraph (f), A' is the supergraph of

$$\hat{f}(\xi) = \sup_v (\xi\cdot v - f(v)).$$

and we obtain

<u>Thm</u>: The is a 1-1 corresp. between conv. l.s.c. $f: V \to \mathbb{R} \cup \{+\infty\}$ such that $f \not\equiv +\infty$, and similar functions $g: V' \to \mathbb{R} \cup \{+\infty\}$ given by

$$g(\xi) = \sup_v \xi\cdot v - f(v) \qquad f(v) = \sup_\xi \xi\cdot v - g(\xi)$$

Now let's return to the original project of a prob measure μ on $\mathbb{R}^n$ (not supported in a hyperplane) and proving that the ~~image~~ of $J \mapsto \bar{x}_J$ is the interior of the convex hull of Supp μ. Suppose to simplify that Supp μ is compact.

The idea will be to form

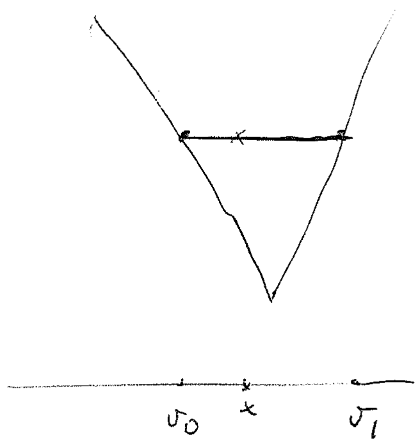
$$W(x) = \sup_\xi \xi\cdot x - \log Z(\xi)$$

and to take an interior point of $\{x \mid W(x) < \infty\}$. Call this point x_0. Let's look at the supergraph of W

What I really want I think is a point (x_0, a_0) contained in the interior of the supergraph. In finite dimensions we can get such a point over an x_0 which is in the interior of a simplex in V of the same dimension whose vertices have finite W. Thus if $v_0, v_1, \dots, v_q$ $\quad q = \dim V$ are affinely independent with $W(v_j) < \infty$, then by convexity

$$W(\Sigma t_j v_j) \leq \sum t_j W(v_j)$$

and so we get an upper bound on W near x_0



Next suppose $W(x) \leq C$ for $|x - x_0| \leq \varepsilon$. Thus

$$* \qquad J(x - x_0) + Jx_0 - \log Z(J) \leq C$$

for all J, all x with $|x - x_0| \leq \varepsilon$.

Might as well suppose $x_0 = 0$, $\varepsilon = 1$.

$$\forall J, \ |x| \leq 1 \implies J \cdot x - \log Z(J) \leq C$$

$$\implies \forall J \quad |J| - \log Z(J) \leq C$$

$$\implies -\log Z(J) \leq C - |J|$$

This implies that $-\log Z(J)$ has a maximum. If J_0 is the maximum value, then $\bar{x}_{J_0} = 0$.

In general given* for all J, all x with $|x - x_0| \leq \varepsilon$ one gets $|J| + Jx_0 - \log Z(J) \leq C$

so $Jx_0 - \log Z(J)$ has a maximum; if it occurs at J_0 then $x_0 = x_{J_0}$.

March 26, 1985

Review convex analysis.

V locally convex Hausdorff top. v.s.

V^* its dual.

$f : V \longrightarrow \mathbb{R} \cup \{+\infty\}$ convex and l.s.c.

Define f^* on V^* by

$$f^*(\xi) = \sup_v \left(\xi \cdot v - f(v)\right)$$

Assertion: Assume $f \not\equiv +\infty$. Then f^* maps V^* to $\mathbb{R} \cup \{+\infty\}$ and it is convex and l.s.c. for the weak $*$ topology on V^*. Moreover

$$f(v) = \sup_\xi \left(\xi \cdot v - f^*(\xi)\right)$$

Proof. Put $\operatorname{epi}(f) = \{(v,a) \in V \times \mathbb{R} \mid a \geq f(v)\}$ epigraph

f l.s.c. $\Longrightarrow$ $\operatorname{epi}(f)$ closed

f convex $\Longrightarrow$ " convex

Lemma: If $(v_0, a_0) \notin \operatorname{epi}(f)$, there exists $(\xi, c) \in V^* \times \mathbb{R}$ ch that $f(v) \geq \xi \cdot v - c \quad \forall v$

$$a_0 < \xi \cdot v_0 - c$$

(This is the non-trivial point: ~~we can separate~~ $\operatorname{epi}(f)$ intersection of all $\operatorname{epi}(\xi \cdot v - c)$ containing it.)

Do rest of the proof. As $\exists\, v$ with $f(v) < \infty$, clearly $f^*(\xi) > -\infty$. As f^* is the sup of linear weak * cont. functions, it is convex and weak * l.s.c.

(Criterion: $f : V \to \mathbb{R} \cup \{+\infty\}$ convex $\iff$ epi(f) convex
l.s.c. $\iff$ closed)

$\forall v, \forall \xi,\ f^*(\xi) \geq \xi \cdot v - f(v) \implies f(v) \geq \sup_\xi (\xi \cdot v - f^*(\xi))$.
To prove equality fix v_0 and let $a_0 < f(v_0)$. By lemma $\exists\, (\xi_0, c_0)$ $\ni$

$$(\forall v,\ f(v) \geq \xi_0 \cdot v - c_0) \iff c_0 \geq f^*(\xi_0)$$

and $\quad a_0 < \xi_0 \cdot v_0 - c_0$. So

$$a_0 < \xi_0 \cdot v_0 - f^*(\xi_0) \leq \sup_\xi (\xi \cdot v - f^*(\xi))$$

~~...~~ Since a_0 is arbitrary $< f(v_0)$, we get

$$f(v_0) \leq \sup_\xi (\xi \cdot v - f^*(\xi))$$

Proof of Lemma: As epi$(f) \subset V \times \mathbb{R}$ is closed convex the Hahn–Banach thm. implies epi(f) is the intersection of the closed affine half spaces containing it. A closed affine half space is ~~described~~ of the form

$$H = \{(v, a) \mid \xi \cdot v - ba \leq c\}$$

with $(\xi, b, c) \in V^* \times \mathbb{R} \times \mathbb{R}$. One can suppose $b = -1, 0, 1$.

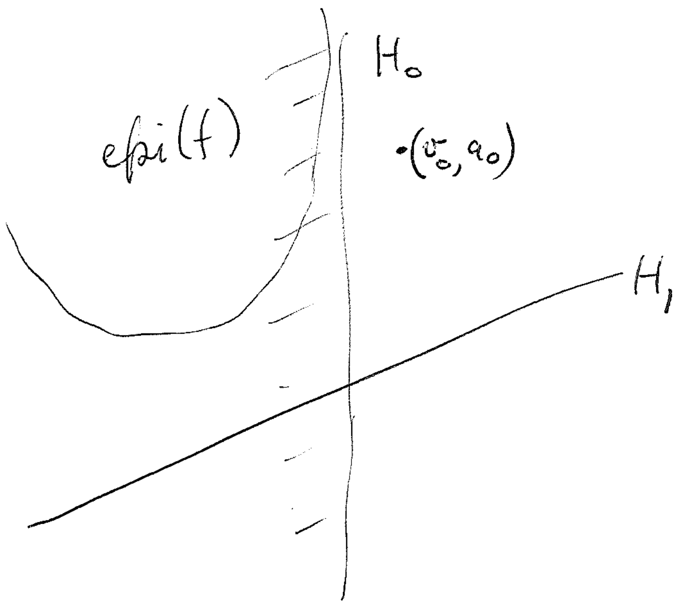
Using the assumption $\exists\, v$ with $f(v) < +\infty$ we see that $H \supset$ epi(f) forces $b = 0$ or 1. If $b = -1$ then H contains (v, a) for $a \geq f(v)$, so

$\xi \cdot v + a \leq c$

for arbitrarily large a which is impossible.

Divide the H containing $\mathrm{epi}(f)$ into $\mathcal{H}^+$, $\mathcal{H}^0$ according as $b=1$ or $b=0$. The lemma says $\mathrm{epi}(f) = \bigcap_{H\in\mathcal{H}^+} H$ and we know $\mathrm{epi}(f) = \bigcap_{H\in\mathcal{H}^+\cup\mathcal{H}^0} H$.

So it is enough to take $(v_0, a_0) \notin H_0$ with $H_0 \in \mathcal{H}^0$ and show $\exists\ H \in \mathcal{H}^+$ with $(v_0, a_0) \notin H$. Picture:



Claim $\mathcal{H}^+ \neq \phi$. $\exists\, v$ with $f(v) < +\infty$, so $\exists\ a \ni (v,a) \notin \mathrm{epi}(f)$, so there is an H_1 containing $\mathrm{epi}(f)$ but not (v,a), and clearly $H_1 \notin \mathcal{H}_0$ so $H_1 \in \mathcal{H}^+$.

Let $H_1 =$ $\mathrm{epi}(\xi_1\cdot v - c_1)$. Now consider $H_\lambda = \mathrm{epi}\left((\lambda\xi_0 + \xi_1)\cdot v - c_1\right)$, where $H_0 = \{(v,a) \mid \xi_0\cdot v \leq 0\}$ and $\lambda > 0$. Then $H_\lambda \supset \mathrm{epi}(f)$.

$(v,a) \in \mathrm{epi}(f) \Rightarrow \xi_0 \cdot v \leq 0$

$\Rightarrow \xi_1\cdot v - c_1 \leq a$

$\Longrightarrow (\lambda\xi_0 + \xi_1)\cdot v - c_1 \leq a$

$\Longrightarrow (v,a) \in \mathcal{H}_\lambda$

On the other hand as $\lambda \longrightarrow +\infty$, $\mathcal{H}_\lambda$ approaches $\mathcal{H}_0$ which doesn't contain (v_0, a_0). So for large λ, $\mathcal{H}_\lambda$ doesn't contain (v_0, a_0). QED.

Remarks: 1) This Hahn-Banach thm. for closed convex sets implies that a closed convex set is weakly closed. So a l.s.c. convex f is l.s.c. for the weak topology.

2) To finish the above one would like to start with a $g : V^* \longrightarrow \mathbb{R} \cup \{+\infty\}$ and define

$$g^*(v) = \sup_{\xi \in V^*} \xi \cdot v - g(\xi)$$

and then recover g as the conjugate of g^*. Thus we need to identify V with the dual of V^*. Question: Is it true that V is the dual of V^* for the $\sigma(V, V^*)$ (or weak star) topology?

Yes. Suppose $\lambda : V^* \longrightarrow \mathbb{R}$ is continuous w.r.t. the $\sigma(V^*, V)$-topology. Then $\exists\ v_1, \dots, v_k \in V$ such that

$$|v_i \cdot \xi| < 1 \qquad i = 1, \dots, k \implies |\lambda(\xi)| < 1.$$

It follows that λ kills $\mathrm{Ker}\{V^* \xrightarrow{(v_i)} \mathbb{R}^k\}$, so λ is a linear combination of the v_i.

So we can identify the transforms f^* on V^* for $f : V \to \mathbb{R} \cup \{+\infty\}$ l.s.c. convex, $f \not\equiv +\infty$ with the set of similar $g : V^* \longrightarrow \mathbb{R} \cup \{+\infty\}$ but where l.s.c. refers to the weak $*$ topology on V^*

April 3, 1985

The problem is the weak law of large numbers in infinite dimensions, in particular, to show that if G is an open set containing $\bar{x}$, then $\mu_n(G) \uparrow 1$. Already this is interesting in the Gaussian case in which μ_n is simply a rescaling of μ; one is interested in ~~showing~~ $\mu(\sqrt{n}\, G) \uparrow 1$, and more specifically the rate of convergence.

Let's first check the scaling. Suppose we have a Gaussian measure on V; say we write formally

$$\mu(dx) = e^{-\frac{1}{2} x \cdot A x}\, dx/\mathcal{N}$$

and then rigorously we have

$$Z(J) = \int e^{J \cdot x} \mu(dx) = e^{-\frac{1}{2} Q(J)}$$

with Q positive definite (formally $Q(J) = J \cdot A^{-1} J$). Then μ_n has the partition function

$$Z(J/n)^n = e^{-\frac{1}{2} n Q(J/n)} = e^{-\frac{1}{2}\frac{1}{n} Q(J)}$$

which formally means

$$\mu_n(dx) = e^{-\frac{1}{2} n\, x \cdot A x}\, dx/\mathcal{N}$$

and so μ_n is μ rescaled under $x \to \sqrt{n}\, x$. Thus

$$\mu_n(G) = \mu(\sqrt{n}\, G)$$

Next I want to take G to be a ball $\{x \mid \|x\| \leq 1\}$ where $\|x\|$ is a norm on V. Actually $\| \ \|$ can be just densely defined; if one shows $\mu(\sqrt{n}\, G) \longrightarrow 1$ as n goes to infinity, then the measure is supported on the Banach space $\bigcup \sqrt{n}\, G$ defined by this norm.

Lets get to the essential example where everything is Gaussian. Suppose $V = \mathbb{R}^{\mathbb{N}_{>0}}$ with

$$\mu(dx) = \prod e^{-\frac{1}{2} a_n x_n^2} \frac{\sqrt{a_n}\, dx_n}{\sqrt{2\pi}}$$

and $\|x\|^2 = \frac{1}{2}\sum_n b_n x_n^2$. Thus I want to compute

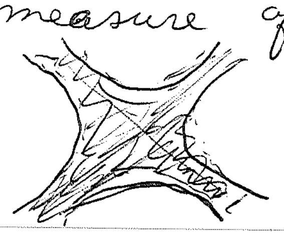
$$\sigma(r) = \int_{\frac{1}{2} x\cdot Bx \leq r} \mu(dr)$$

It's simpler to compute the Laplace transform

$$\int e^{-sr}\, d\sigma(r) = \int e^{-s\frac{1}{2} x\cdot Bx} \mu(dx)$$

$$= \det\left(\frac{A+sB}{A}\right)^{-1/2} = \prod \left(1 + s\frac{b_n}{a_n}\right)^{-1/2}$$

(Incidently this same calculation says something when B isn't positive definite, i.e. we are computing the measure of a hyperbolic ~~sphere~~ ball



)

Now let's use the estimates involved in the 1-dim'l Cramer thm.

$$\int_0^\infty e^{sr}\, d\sigma(r) = \prod\left(1-s\frac{b_n}{a_n}\right)^{-1/2}$$

[redacted]

$$\geq \int_{[R,\infty)} e^{sr}\, d\sigma(r) \overset{s\geq 0}{\geq} \int_{[R,\infty)} e^{sR}\, d\sigma(r) = e^{sR}\left(1-\sigma(R)\right)$$

Thus

(*) $$1-\sigma(R) \leq e^{-sR}\left[\prod\left(1-s\frac{b_n}{a_n}\right)\right]^{-1/2}$$

assuming $s \geq 0$.

The first condition is

1) $$\sum \frac{b_n}{a_n} < \infty .$$

This [redacted] implies that the infinite product converges for small s. [redacted] It is necessary for the ball $\frac{1}{2} xBx \leq r$ $(r>0)$ to have positive measures. Assume 1) holds.

Then the infinite product in (*) converges for all s, so we can use (*) for all $s \geq 0$ such that

$$1 - s\frac{b_n}{a_n} > 0 \quad \text{or} \quad s < \inf\left(\frac{a_n}{b_n}\right)$$

Thus we have exponential convergence

$$1-\sigma(R) \leq e^{-sR} \text{ const}$$

for any $0 \leq s < \inf\left(\frac{a_n}{b_n}\right)$.

So $\overline{\lim} \frac{1}{R} \log (1 - v(R)) \leq -s \quad \forall s \ni 0 \leq s < \inf\left(\frac{a_k}{b_k}\right)$

or

$$\overline{\lim} \frac{1}{R} \log (1 - v(R)) \leq -\inf\left(\frac{a_k}{b_k}\right)$$

Note this is exactly the Cramer estimate. If $R = n$

$$v(n) = \mu\left(\tfrac{1}{2} xBx \leq n\right) = \mu\left(\sqrt{n} \cdot \{x \mid \tfrac{1}{2} xBx \leq 1\}\right)$$
$$= \mu_n\left(\{x \mid \tfrac{1}{2} xBx \leq 1\}\right)$$

$$1 - v(n) = \mu_n\left(\{x \mid \tfrac{1}{2} xBx > 1\}\right)$$

$W(x) = \frac{1}{2} \sum a_k x_k^2$; we want the inf of this over $\{x \mid \frac{1}{2} \sum b_k x_k^2 \geq 1\}$.
Rescale $x_k \mapsto \sqrt{2/b_k}$; we want

$$\inf\left\{ \sum \frac{a_k}{b_k} x_k^2 \;\Big|\; \sum x_k^2 \geq 1 \right\}$$

and this is clearly $\inf\left(\frac{a_k}{b_k}\right)$.

~~Let~~ let's consider a (centered) Gaussian prob measure μ on a top vector space V (subject to some later hypotheses). Gaussian means the image under any linear map $V \to R$ is Gaussian and this should be the same as

$$Z(J) = \int e^{J \cdot x} \mu(dx) = e^{Q(J)}$$

where Q is a non-negative quadratic form on V^*

Let's consider the standard form for such a measure. The fn. Q is equivalent to a map $G: (V^*) \longrightarrow (V^*)^*$ which is symmetric: $G^t = G$. We suppose G

continuous and V reflexive.

Better, we suppose Q is given by a continuous linear map $G: V^* \longrightarrow V$ which is symmetric $G^t = G$ and injective with dense image. (One knows $Q(J)$ is convex, quadratic, and l.s.c. Perhaps one could show Q continuous.) Formula

$$Q(J) = \tfrac{1}{2} J \cdot GJ.$$

Continue with the standard picture: We use Q to define an inner product on V^* and form the completion H of V^* with respect to this inner product. We let $W \subset V$ be the subspace of $x \in V$ such that the associated linear fnl on V^* is continuous for the inner product. W can be identified with H

The inner product is

$$\langle J, J' \rangle = J \cdot GJ'$$

So let us try to describe the picture. One has maps where H is a Hilbert space

$$V^* \hookrightarrow H \hookrightarrow V$$

which are embeddings with dense image such that the first and second are transposes of each other (assuming V reflexive). The composition is this map G.

Next try to compute the transform of $Q(J) = \frac{1}{2} J \cdot GJ$.

$$W(x) = \sup_J \left(J \cdot x - \tfrac{1}{2} J \cdot GJ \right)$$

If $x = GJ_0$, then $J = J_0$ should be the critical point and we would get $\quad W(x) = \frac{1}{2} J_0 \cdot GJ_0 = \frac{1}{2} x \cdot G^{-1} x$

But actually $W(x)$ should be finite for $x \in H$, and then $W(x)$ should be a natural extension of $\frac{1}{2} x \cdot G^{-1} x$.

April 5, 1985

Let's consider a Gaussian discrete random process and analyze what it means for it to be a martingale.

Let (Ω, μ) be a probability space, let $f_1, f_2, \dots$ be a sequence of random variables on Ω (real-valued). One says $\{f_n\}$ is a Gaussian process if the map

$$(f_1, \dots, f_n) : \Omega \longrightarrow \mathbb{R}^n$$

pushes μ forward to a Gaussian measure μ_n on $\mathbb{R}^n$, for each n. (Assume centered at 0.)

One might as well suppose $\Omega \subset \mathbb{R}^\infty$ and the f_n are the coordinate functions x_n.

Consider $\mathcal{H} = L^2(\Omega, \mu)$, and let $V \subset \mathcal{H}$ be the closed subspace spanned by the $\{f_n\}$. I think we know that if $\mu_n = (f_1, \dots, f_n)_* \mu$, then

$$L^2(\Omega, \mu) = \varinjlim L^2(\mathbb{R}^n, \mu_n)$$

and the space of polynomial functions $S(V)$ is dense.

Assertion: a discrete Gaussian random process $\{f_n\}$ is equivalent to a Hilbert space V together with a sequence $f_1, \dots, f_n, \dots$ in V spanning V. Equivalently a non-negative symmetric matrix $\langle f_i f_j \rangle$.

Next recall that for Gaussian r.v.'s independence is equivalent to orthogonality. $\{f_n\}$ will be an L^2 martingale when $f_n - f_{n-1}$ is orthogonal to $L^2(\mathbb{R}^{n-1}, \mu_{n-1})$.

It should be enough that $f_n - f_{n-1} \perp V_{n-1} = \langle f_1, \dots, f_{n-1} \rangle$ (think of the Wick process)

Conclude: A Gaussian martingale is of the form

$$f_n = g_1 + g_2 + \cdots + g_n$$

where the g_n are a sequence of independent Gaussian r.v.'s.

The martingale convergence theorem says that such a sequence $f_n = g_1 + \cdots + g_n$ converges almost everywhere.

Let's now consider the Hilbert space $V = \ell^2$ which we identify with a space of functions on $\mathbb{R}^\infty$. Better: I want to take the standard Gaussian measure $e^{-x^2/2}\, dx/\sqrt{2\pi}$ on $\mathbb{R}$ which describes a Gaussian r.v. with standard deviation 1. Then take the product measure

$$\mu = \prod_{n=1}^{\infty} e^{-\frac{1}{2}x_n^2} \frac{dx_n}{\sqrt{2\pi}} \qquad \text{on } \mathbb{R}^\infty.$$

We want $\Omega \subset \mathbb{R}^\infty$ to be subspace carrying μ. The idea is that Ω should be a Banach space or top v.s. such that there is an embedding $\ell^2 \subset \Omega$ carrying Gaussian cylinder measure on ℓ^2 to a measure on Ω.

Now take $g_n = b_n x_n$ where x_n is the nth coordinate function and $f_n = \sum_{k \le n} b_k x_k$. This is a Gaussian martingale, so we see that provided

$\sum b_k^2 < \infty$ the series $\sum_j b_k x_k$ converges almost everywhere on Ω.

April 6, 1985

I want to now try to prove that if I have a sequence of random variables $g_1, g_2, \ldots$ which are independent and of mean zero and such that $\sum \langle g_k^2 \rangle < \infty$, then the series $\sum g_k$ converges almost everywhere.

Put $f_n = g_1 + \cdots + g_n$. We want to prove this sequence converges a.e. So we look at those x such that $f_n(x)$ fails to converge. If the sequence $f_n(x)$ fails to converge we know it is unbounded or has at least 2 limit points. If it has at least two limit points, then there is an interval (a, b) with rational endpoints such that $f_n(x)$ is infinitely often below a and above b. Let

$$S_{(a,b)} = \left\{ x \,\middle|\, \begin{array}{ll} f_n(x) \geq b & \text{i.o.} \\ \text{and } f_n(x) \leq a & \text{i.o.} \end{array} \right\}$$

Then the subset of x for which $f_n(x)$ doesn't converge is

$$\bigcup_{\substack{a,b \in \mathbb{Q} \\ a<b}} S_{(a,b)} \cup \underbrace{\{x \mid f_n(x) \to +\infty\}}_{S_{+\infty}} \cup \underbrace{\{x \mid f_n(x) \to -\infty\}}_{S_{-\infty}}$$

It suffices to show $S_{(a,b)}$ and $S_{+\infty}$, $S_{-\infty}$ are null-sets.

The key point in the proof will be to estimate the probability that with k fixed one has $f_n - f_k = g_{k+1} + \cdots + g_n \geq \varepsilon$ for some n.

~~Better what I want to know is how likely it is for~~

~~[illegible]~~ Better: Given k, find

$$P\{x \mid \exists n \text{ with } (g_{k+1}+\dots+g_n)(x) \geq \varepsilon\}$$

Let's analyze this for $k=0$. We are interested in the set where $f_n \geq \varepsilon$ for some n. We decompose this into

$$B_k : \quad f_1, \dots, f_{k-1} < \varepsilon,\ f_k \geq \varepsilon$$

Then for $n \geq k$

$$\int_{B_k} f_n = \int_{B_k} f_k \geq \varepsilon P(B_k)$$

Cauchy-Schwartz + $\int_B f = \int \chi_B \cdot f$

so
$$\varepsilon P\Big(\bigcup_{k\leq n} B_k\Big) \leq \int_{\bigcup_{k\leq n} B_k} f_n \leq P\Big(\bigcup_{k\leq n} B_k\Big)^{1/2} \|f_n\|$$

$$\varepsilon P\Big(\bigcup_{k\leq n} B_k\Big)^{1/2} \leq \|f_n\|$$

$$P\Big(\bigcup_{k\leq n} B_k\Big) \leq \frac{1}{\varepsilon^2}\|f_n\|^2 \qquad (\varepsilon > 0 \text{ needed})$$

$$P\Big(\bigcup_{k} B_k\Big) \leq \frac{1}{\varepsilon^2}\|f\|^2 \qquad f = \lim f_n \text{ in } L^2$$

(Recall Chebyshev:

$$\|f\|^2 \geq \int_{|f|\geq\varepsilon} f^2 \geq \varepsilon^2 P\{|f| \geq \varepsilon\} \implies P\{|f|\geq\varepsilon\} \leq \frac{1}{\varepsilon^2}\|f\|^2$$

so this is the same kind of inequality.)

Thus we have the estimate

$$\boxed{\forall \varepsilon > 0 \quad P\{\exists n \text{ with } (g_{k+1} + \cdots + g_n) \geq \varepsilon\} \leq \frac{1}{\varepsilon^2} \sum_{j=k+1}^{\infty} \|g_j\|^2}$$

Now we should be able to show that the set where $f_n \longrightarrow +\infty$ is a null set. But if $f_n(x) \longrightarrow +\infty$, then for any ε one has $f_n(x) \geq \varepsilon$ for n large enough. Thus $\forall \varepsilon$

$$S_\infty = \{x \mid f_n(x) \longrightarrow +\infty\} \subset \{x \mid \exists n \quad f_n(x) \geq \varepsilon\}$$

so $\quad P(S_\infty) \leq \dfrac{1}{\varepsilon^2} \displaystyle\sum_1^\infty \|g_j\|^2 \quad$ for all ε, etc.

Next let us take up bounding $P(S_{a,b})$. Here the idea is to note

$$S_{a,b} = \bigcap_r S^r_{a,b}$$

where $S^r_{a,b}$ is the set of x such that we have $\geq r$ r crossings. Thus S^1 means $f_k(x) \leq a$ for some k, and S^2 means $\exists k, \ell \ni k < \ell$ $f_k(x) \leq a$ and $f_\ell(x) \geq b$, etc.

Introduce stopping times

$$N_1(x) = \inf\{k \mid f_k(x) \leq a\}$$

$$N_2(x) = \inf\{\ell \mid \ell > N_1(x) \text{ and } f_\ell(x) \geq b\}$$

Then $S^r_{ab} = \{x \mid N_r(x) < \infty\}$. Now what we have to do is estimate $P(S^{r+1})$ in terms of $P(S^r)$.

First observe that S^{r+1} decomposes according to the values of $N_1, \ldots, N_{r+1}$. Similarly for S^r. Let's fix the values of $N_1, \dots, N_r$, i.e. we take one of the components of S^r, denote it $S^r(k_1, \dots, k_r)$.

$$S^r(k_1, \dots, k_r): \quad f_1 \cdots f_{k_1 - 1} > a, \ f_{k_1} \leq a$$
$$f_{k_1+1}, \dots, f_{k_2-1} < b, \ f_{k_2} \geq b$$
$$\vdots$$

Next we want the part of S^{r+1} which is contained in $S^r(k_1, \dots, k_r)$. This is

$$S^{r+1} \cap S^r(k_1, \dots, k_r) = \coprod_{\ell > k_r} S^{r+1}(k_1, \dots, k_r, \ell).$$

We look at a point x) of this intersection, and note that the sequence

$$f_{k_r+1}(x), \ f_{k_r+2}(x), \ \dots$$

has a jump of at least $b-a$.

Now I want to use the independence. $S^r(k_1, \dots, k_r)$ is a subset coming from Ω_{k_r} which means that it is equivalent to the product of a subset of Ω_{k_r} and the space of sequences (g_j), $j > k_r$. It seems that $S^{r+1} \cap S^r(k_1, \dots, k_r)$ is the product of this subset of Ω_{k_r} and the space of sequences such that $g_{k_r+1} + \dots + g_j$ has a jump of $\varepsilon = |b-a|$ of the right sign. Thus we should get the ~~estimate~~

$$P(S^{r+1} \cap S^r(k_1, \dots, k_r)) \leq P(S^r(k_1, \dots, k_r)) \frac{1}{\varepsilon^2} \sum_{j > k_r} \|g_j\|_2^2$$

and finally because $k_r \geq r$ we get

$$P(S^{r+1} \cap S^r(k_1,\dots,k_r)) \leq P(S^r(k_1,\dots,k_r)) \cdot \frac{1}{\varepsilon^2} \sum_{j>r} \|g_j\|_2^2$$

which we can add up to get

$$P(S^{r+1}) \leq P(S^r) \cdot \underbrace{\frac{1}{\varepsilon^2} \sum_{j>r} \|g_j\|_2^2}$$

Then for large r the factor is < 1 so one sees that $P(S^r) \downarrow 0$.

April 8, 1985

Goal: Proof of the martingale a.e. convergence thm.

Recall the definition: The probability space Ω is assumed to be the limit of an inverse system

$$\cdots \longrightarrow \Omega_n \xrightarrow{\pi^n_{n-1}} \Omega_{n-1} \longrightarrow \cdots \longrightarrow \Omega_1$$

and the measure μ on Ω induces μ_n on Ω_n. $\{f_n\}$ is a sequence of random variables on Ω such that f_n comes from $\bar{f}_n$ on Ω_n. This sequence is a martingale when

$$\begin{cases} f_n \in L^1 & \text{and} \\ (\pi^n_k)_* \bar{f}_n = \bar{f}_k & n \geq k \end{cases}$$

Here π_* denotes conditional expectation:

$$\int_A \pi_*(f) = \int_{\pi^{-1}(A)} f$$

It is defined by virtue of Radon-Nikodym.

Consequence:
$$\left| \int_A \pi_*(f) \right| \leq \int_{\pi^{-1}(A)} |f| = \int_A \pi_* |f|$$

holds for all A $\Rightarrow$
$$|\pi_*(f)| \leq \pi_* |f|.$$

So for a martingale

$$|\bar{f}_k| = |(\pi^n_k)_*(\bar{f}_n)| \leq (\pi^n_k)_* |\bar{f}_n| \implies |\bar{f}_k|_1 \leq |\bar{f}_n|_1 \quad \text{for } k \leq n$$

The martingale convergence thm. says that if $|f_n|_1$ is bounded, then f_n converges almost everywhere.

Recall the concept of stopping time. This is a meas. fn: $N : \Omega \longrightarrow \mathbb{N}_{>0}$ such that for each n $\{x \mid N(x) \leq n\}$ comes from Ω_n. What this gives us is a decomposition of Ω

$$\Omega = \coprod_{n=1}^{\infty} B_n \qquad\qquad B_n : N = n$$

where B_n comes from Ω_n.

~~[illegible]~~ A stopping time N allows us to construct two martingales associated to a given martingale. The first represents running the process to the stopping time and then fixing the value at this point. This is

$$n \longmapsto f_{n \wedge N}, \quad \text{i.e.} \quad f_{n\wedge N}(x) = \begin{cases} f_n(x) & n \leq N(x) \\ f_{N(x)}(x) & n \geq N(x) \end{cases}$$

The other martingale is defined on $\{x \mid N(x) < \infty\}$ assuming this set has positive probability and is

$$n \longmapsto f_{n+N} \qquad \text{i.e.} \quad f_{n+N}(x) = f_{n+N(x)}(x).$$

We can describe this process as being f_{n+k} over B_k

(One probably has to write out some details about how the Ω's are defined for this process. However it is clear that the basic averaging property holds.)

Remarks: From the Hilbert space viewpoint one has an increasing sequence of von Neumann algebras

$$L^\infty(\Omega_1) \subset L^\infty(\Omega_2) \subset \cdots\cdots$$

and the measure gives compatible traces. Conditional expectation is then defined to be adjoint to inclusion relative to the trace pairing. A stopping time is an operator $N = \sum_{1\le n\le\infty} n E_n$ such that $E_n \in L^\infty(\Omega_n)$.

Let go back to the martingale $\{f_n\}$ and the stopping time N. I assume we have an L^2-martingale i.e. each f_n is in L^2 and $\|f_n\|_2$ is bounded. In this case we know f_n converges to f in $L^2(\Omega)$, because the differences $f_n - f_{n-1}$ are mutually orthogonal.

From now on I want to think of this setup as being determined by the limit function f.

Now consider the stopping time N and the family $f_{n\wedge N}$. I want to see clearly that this is a martingale. To do this we decompose everything according to the values of N. Look at B_k: $N=k$. At a point x of B_k (i.e. $N(x)=k$) we have

$$f_{n\wedge N}(x) = \begin{cases} f_n(x) & n \le k \\ f_k(x) & n \ge k \end{cases}$$

Suppose I introduce the orthogonal sequence $g_j = f_j - f_{j-1}$. Then $f_{n\wedge N}$ is over B_k the sequence of partial sums of the series

$$\sum_j \begin{cases} g_j & j \le k \\ 0 & j > k \end{cases}$$

Let's carefully discuss the decomposition of L^2 that a stopping time gives.

Ω_1 decomposes into $\pi_1 B_1$ and its complement

Ω_2 ———————— $\pi_2 B_1$, $\pi_2 B_2$, and the rest

Picture: We have this sequence of von Neumann algebras

$$(*) \qquad L^\infty(\Omega_1) \subset L^\infty(\Omega_2) \subset \cdots$$

with limit $L^\infty(\Omega)$ and we have the measure on Ω giving rise to traces on the $L^\infty(\Omega_n)$ and the conditional expectation maps

$$L^1(\Omega_1) \longleftarrow L^1(\Omega_2) \longleftarrow \cdots\cdots\cdots$$

Now any L^2-martingale associated to the filtration $(*)$ is equivalent to an $f \in L^2(\Omega)$; then f_n is the orthogonal projection of f onto $L^2(\Omega_n)$.

The point is that any compatible family of $f_n \in L^1(\Omega_n)$ converges almost everywhere provided $\|f_n\|_1$ are bounded.

I would like to write out the proof at least in the case where $f_n \in L^2(\Omega_n)$. We fix an interval (a, b) and let S be the subset of Ω where $f_n \leq a$ for inf. many n and $f_n \geq b$ for infinitely many n. We must show S has measure zero.

We define stopping times

N_1 = first n such that $f_n \leq a$ or $+\infty$

N_2 = first $n > N_1$ ——— $f_n \geq b$ or $+\infty$

Better

$$N_1(x) = \inf\{n \mid f_n(x) \geq a\}$$
$$N_2(x) = \inf\{n \mid n > N_1(x) \text{ and } f_n(x) \geq b\}$$
$$N_3(x) = \inf\{n \mid n > N_2(x) \text{ and } f_n(x) \leq a\} \qquad \text{etc.}$$

Then if $S_r = \{x \mid N_r(x) < \infty\}$ we have $\bigcap S_r = S$, so we want to show $P(S_r) \to 0$.

Actually I want to first estimate $P\{N_2 < \infty\}$. On $N_2 < \infty$ we have $f_{N_2} \geq b$, $f_{N_1} \leq a$ so

$$\int_{N_2<\infty} f_{N_2} - f_{N_1} \geq (b-a)\,P\{N_2 < \infty\}$$

On the other hand

$$\int_{N_2<\infty} f_{N_2} - f_{N_1} = \int_{N_2<\infty} f - f_{N_1} \leq P\{N_2 < \infty\}^{1/2} \cdot \|f - f_{N_1}\|_2$$

so putting these together yields

$$P\{N_2 < \infty\} \leq \frac{1}{(b-a)^2} \|f - f_{N_1}\|_2^2$$

It seems the same argument shows more generally

$$P\{N_2 < \infty\} \leq \frac{1}{(b-a)^2} \|f_{N_2} - f_{N_1}\|_2^2 .$$

A similar argument should yield

$$P\{N_k < \infty\} \leq \frac{1}{(b-a)^2} \|f_{N_k} - f_{N_{k-1}}\|^2$$

and on the other hand

$$\|f_{N_k} - f_{N_{k-1}}\|^2 \leq \|f - f_{k-1}\|^2$$

which goes to zero as $k \to \infty$.

Now I want to check that the above is correct. I am using that for an L^2-martingale and stopping time $N_2 > N_1$ that

$$f = (f - f_{N_2}) + (f_{N_2} - f_{N_1}) + f_{N_1}$$

is an orthogonal direct sum. We want to check this sort of thing very carefully.

We begin with the increasing sequence

$$L^2(\Omega_1) \subset L^2(\Omega_2) \subset \cdots\cdots$$

and the orthogonal projections backward. Recall that if N is a stopping time, then $N : \Omega \to \mathbb{N}_{>0} \cup \{+\infty\}$ and $\{x \mid N(x) = n\}$ comes from Ω_n. We define for $f \in L^2(\Omega)$

$$(f_N)(x) = \begin{cases} f_{N(x)}(x) & \text{if } N(x) < \infty \\ f(x) & \text{if } N(x) = \infty \end{cases}$$

Put $B_n = \{x \mid N(x) = n\}$. Now I want to describe f_N more precisely with the goal of showing $f - f_N$ and f_N are orthogonal. Over B_∞ this is clear. First we should point out that we propose to compute as follows:

$$\int (f - f_N) f_N = \sum_n \int_{B_n} (f - f_N) f_N = \sum_n \int_{B_n} (f - f_n) f_n$$

But by assumption B_n comes from Ω_n and

$$\int f f_n \chi_{B_n} = \int (\pi_n)_* f \cdot f_n \chi_{B_n} = \int f_n^2 \chi_{B_n}$$

so it's clear that $\int (f - f_N) f_N = 0$.

Similarly if $N_1 \leq N_2$ are stopping times, then $f_{N_2} - f_{N_1} \perp$ anything occurring before N_1.

<u>Proof</u>. Decompose according to the values of N_1; it's enough to assume N_1 is constant say $N_1 = k$ and then we want to prove that projecting onto $L^2(\Omega_k)$ sends $f_{N_2} - f_k$ into zero. But this should be the same as projecting $f - f_k$

which we know goes to zero.

To really be convincing I ought to assign to N a quotient Ω_N of Ω such that
i) f_N is the orthogonal projection ~~of~~ f on $L^2(\Omega_N)$
ii) if $N \geq N'$ then $L^2(\Omega_{N'}) \subset L^2(\Omega_N)$.

Let $\Omega = B_1 \sqcup \cdots \sqcup B_\infty$ be the decomposition by the values of N. We want to define a quotient tower

$$\begin{tikzcd}
\Omega_1 = & B_{11} & \sqcup & \Omega_1 - B_{11} & & & & \\
\Omega_2 = & B_{12} \arrow[u] & \sqcup & B_{22} \arrow[u] & \sqcup & \Omega_2 - (\quad) \arrow[ull] & & \\
\Omega_3 = & B_{13} \arrow[u] & \sqcup & B_{23} \arrow[u] & \sqcup & B_{33} & \sqcup & \text{comp.}
\end{tikzcd}$$

Let

$$\overline{\Omega}_1 = B_{11} \sqcup \Omega_1 - B_{11}$$
$$\overline{\Omega}_2 = B_{11} \sqcup B_{22} \sqcup \Omega_2 - (\quad)$$

etc. Thus

$$\Omega_N = \lim \overline{\Omega}_j = B_{11} \sqcup B_{22} \sqcup \cdots$$

Now corresponding to the splitting of Ω into $B_1 \sqcup \cdots \sqcup B_\infty$ one has a decomposition of f into $f_1 + \cdots + f_\infty$. The projection of f onto $L^2(\Omega_N)$ should be $\bar{f}_1 + \cdots + f_\infty$ where $\bar{f}_i$ is the projection of $L^2(B_i)$ onto $L^2(B_{ii})$. To compare this with f_N. Over B_i, f_N is f_i and it should be OK.

Still not clear.

April 9, 1985

Let us begin with our tower

$$\Omega_1 \twoheadleftarrow \Omega_2 \twoheadleftarrow \Omega_3 \twoheadleftarrow \cdots\cdots\cdots \leftarrow \Omega_\infty = \Omega$$

and let the σ-fields of measurable subsets be

$$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3 \subset \cdots\cdots \qquad \mathcal{F}_\infty = \mathcal{F}$$

($\mathcal{F}_n$ = Boolean algebra of projectors in $L^\infty(\Omega_n)$ up to null set equivalence.)

A stopping time N is a map $N: \Omega \longrightarrow \{1 \leq n \leq \infty\}$ such that $\{N \leq n\} \in \mathcal{F}_n$ for each n. Such an N is the same thing as a decomposition

$$\Omega = B_1 \amalg \cdots\cdots \amalg B_\infty \qquad \text{with } B_n \in \mathcal{F}_n.$$

Let's define $\mathcal{F}_N$ to be the set of all $A \in \mathcal{F}$ such that $A \cap B_n \in \mathcal{F}_n$ for all n. $\mathcal{F}_N$ is a σ-algebra. (This means that $\mathcal{F}_N$ is a Boolean algebra - closed under $\cap, \cup, -$ in $\mathcal{F}$ and also countable unions.)

Next note that

$$A \cap \{N \leq n\} = \coprod_{k \leq n} A \cap B_k$$

so $\forall n \left(A \cap B_n \in \mathcal{F}_n\right) \Longrightarrow \forall n \left(A \cap \{N \leq n\} \in \mathcal{F}_n\right)$. And as

$$A \cap B_n = A \cap \{N \leq n\} - A \cap \{N \leq n-1\}$$

the converse is true.

Thus

$$\mathcal{F}_N = \{A \in \mathcal{F} \mid \forall n \ \ A \cap \{N \leq n\} \in \mathcal{F}_n\}$$

If $N' \leq N$, then $A \cap \{N \leq n\} = A \cap \{N' \leq n\} \cap \{N \leq n\}$ so $A \in \mathcal{F}_{N'}$ implies $A \cap \{N' \leq n\} \in \mathcal{F}_n$, and as $\{N \leq n\} \in \mathcal{F}_n$, one has $A \cap \{N \leq n\} \in \mathcal{F}_n$ for all n, so $A \in \mathcal{F}_N$.

Finally what stopping time gives $\mathcal{F}_n$? It is $N \equiv n$, so that $B_n = \Omega$ and $A \in \mathcal{F}_N \Longleftrightarrow A \in \mathcal{F}_n$.

<u>April 12, 1985</u>

Discrete Parameter Martingales, J. Neveu, North-Holland (1975).

<u>Relation with Markov chains.</u> State space E, the probability space is $\Omega = E^{\mathbb{N}}$, the σ-field $\mathcal{B}$ is generated by $X_n : \Omega \to E$, the σ-subfield $\mathcal{B}_p$ is generated by $X_0, \ldots, X_p$. For each $x \in E$ one has P_x on $(\Omega, \mathcal{B})$ given by

$$P_x\left(X_0 = x_0, \ldots, X_p = x_p\right) = \delta_x(x_0)\, P(x_0, x_1) \cdots P(x_{p-1}, x_p)$$

where $P(x,y)$ is a **Markov** matrix (E supposed discrete, i.e. a countable set; $P(x,\cdot)$ is a prob. measure on E for each x). We have a <u>stationary</u> Markov chain

Given $f : E \to \mathbb{R}$, consider f_n on Ω given by $f_n = f(X_n)$. The conditional expectation of f_n rel. to $\mathcal{B}_{n-1}$ is

$$(\pi_{n-1} f_n)(x_0, \ldots, x_{n-1}) = \sum_{x_n} \frac{p(x_0, \ldots, x_n) f(x_n)}{p(x_0, \ldots, x_{n-1})}$$

$$= \sum_{x_n} P(x_{n-1}, x_n) f(x_n) = (Pf)_{n-1}(x_0, \ldots, x_{n-1})$$

In general use $P(x,y)$ to define an operator on fns. on E by

$$(Pf)(x) = \sum_y P(x,y) f(y).$$

Thus

$$f_n = f(X_n) \text{ is a martingale} \iff Pf = f \qquad (f \text{ harmonic})$$

$$\text{——— submartingale} \iff Pf \geq f \qquad (f \text{ subharm.})$$

General definitions. Given $(\Omega, \mathcal{B}, P)$ and $\subset \mathcal{B}_n \subset \cdots \subset \mathcal{B}$.
A process $\{f_n\}$, i.e. sequence of r.r.v. (real random vbls), is called <u>adapted</u> if f_n is meas. rel $\mathcal{B}_n$, and <u>predictable</u> if f_n is meas. rel $\mathcal{B}_{n-1}$. (convention $f_0 = 0$ for predictable)

<u>Doob decomp.</u> Any adapted process $\{f_n\}$ is uniquely a sum $m_n + a_n$ where m_n is a martingale and a_n is predictable.

Why:

$$f_n = m_n + a_n$$

$$\left.\begin{aligned} \pi_{n-1} f_n &= m_{n-1} + a_n \\ f_{n-1} &= m_{n-1} + a_{n-1} \end{aligned}\right\} \Rightarrow \quad a_n - a_{n-1} = \pi_{n-1} f_n - f_{n-1}$$

This tells you how to define the a_n inductively, starting with $a_0 = 0$

Note: $\{f_n\}$ submartingale $\Rightarrow$ $a_n \geq a_{n-1}$, so the predictable process is increasing.

<u>Integration wrt a martingale</u> (also called martingale transform)
Let $\{f_n\}$ be a martingale, let $\{g_n\}$ be predictable. Then

$$h_n = \sum_{k=1}^{n} g_k (f_k - f_{k-1})$$

is a **martingale**: $\pi_{n-1}(h_n) = h_{n-1} + \underbrace{\pi_{n-1}(g_n(f_n - f_{n-1}))}_{g_n \pi_{n-1}(f_n - f_{n-1}) \,=\, 0}$

(hypotheses: $g_k \in L^\infty$, $f_k \in L^1$ or maybe L^q and L^p?)

Example: This is a discrete version of

$$h(t) = \int_0^t g(\omega, s)\, d\beta(s)$$

<u>Quadratic variation</u> of a sequence f_n, $n \in \mathbb{N}$ is

$$V = f_0^2 + \sum_1^\infty |f_k - f_{k-1}|^2$$

More generally we are interested in the sequence

$$V_n = f_0^2 + \sum_{1 \le k \le n} |f_k - f_{k-1}|^2$$

If f_n is a L^2-martingale, then f_n^2 is a submartingale (Cauchy-Schwartz $\Rightarrow$ $|\int f|^2 \le \int |f|^2$ for a prob. measure - use this for conditional exp. which is roughly a family of prob. measures on the fibres)

f_n^2 has a Doob decomposition $f_n^2 = m_n + a_n$ where a_n is increasing predictable.

$$a_n - a_{n-1} = \Pi_{n-1}(f_n^2) - f_{n-1}^2$$

Also $$\Pi_{n-1}(f_n - f_{n-1})^2 = \Pi_{n-1}(f_n^2) - f_{n-1}^2 \qquad \left(E((x-\bar{x})^2) = E(x^2) - \bar{x}^2\right)$$

So $$a_n = \sum_1^n \Pi_{k-1}(f_k - f_{k-1})^2 .$$

If $f_k - f_{k-1}$ are independent r.v. then $\Pi_{k-1}(f_k - f_{k-1})^2$ is a constant, namely, the variance of this variable. In this case the predictable process is just an the increasing sequence of constants $a_n = \|f_n\|^2$

We need the quadratic variation sequence to control the integral.

$$h_n = \sum_{1 \le k \le n} g_k (f_k - f_{k-1}) \qquad \begin{array}{l} f \text{ mart} \\ g \text{ pred.} \end{array}$$

then $h_n - h_{n-1} \perp h_{n-1}$ so

$$\|h_n\|^2 - \|h_{n-1}\|^2 = \int g_k^2 (f_k - f_{k-1})^2$$

Better
$$\Pi_{n-1} h_n^2 - h_{n-1}^2 = \Pi_{n-1}(h_n - h_{n-1})^2$$
$$= \Pi_{n-1}\, g_n^2 (f_n - f_{n-1})^2$$
$$= g_n^2\ \Pi_{n-1}(f_n - f_{n-1})^2 = g_n^2\,(a_n - a_{n-1})$$

Thus
$$\|h\|^2 = \sum_n \int g_n^2\,(a_n - a_{n-1}).$$

Snell's problem: Let Z_n $n \in \mathbb{Z}$ be a sequence of int. r.r.v.'s representing winnings of a gambler. For any stopping time ν we get an expectation $E(Z_\nu)$. The problem is to find the optimal stopping time, i.e. which realizes $\sup_\nu E(Z_\nu)$.

Analogy: Suppose z_n is a sequence in $\mathbb{R}$ and we want to find p with $z_p = \sup z_n$. Then we introduce $x_p = \sup_{n \geq p} z_n$ which is a decreasing sequence and look for the first p with $x_p = z_p$. If this p is $< \infty$, one wins.

In general one sets
$$X_n = \operatorname*{ess\,sup}_{\Lambda_n} \Pi_n(Z_\nu)$$

$\nu \in \Lambda_n$ means $n \leq \nu$ a.e., and $E(Z_\nu^-) < \infty$.

Snell's thm. says that if $\nu_0 = \inf\{n \mid X_n = Z_n\}$ then there is an optimal $\nu \iff \nu_0 < \infty$ a.e. Also $\nu_\varepsilon = \inf\{n \mid X_n < Z_n + \varepsilon\}$ is finite a.e. and satisfies $E(Z_{\nu_\varepsilon}) + \varepsilon \geq \sup_\nu E(Z_\nu)$.

April 14, 1985

Positive supermartingales and potentials

$X_n \geq 0$ for $n \geq 0$; X_n supermart: $\pi_{n-1} X_n \leq X_{n-1}$. This is an analogue of a positive decreasing sequence. The basic convergence thm. implies X_n converges a.e. to X_∞ and $\pi_n X_\infty \leq X_n$ for all n.

Next one has the Doob decomp.

$$X_n = M_n - A_n \qquad A_0 = 0$$

$$\pi_{n-1} X_n = M_{n-1} - A_n$$

$$X_{n-1} = M_{n-1} - A_{n-1} \qquad A_n - A_{n-1} = X_{n-1} - \pi_{n-1} X_n \geq 0$$

so A_n is an increasing predictable process. Put $A_\infty = \sup A_n$. Then

$$\pi_0 A_n \leq \pi_0 A_n + \pi_0 X_n = \pi_0 M_n = X_0$$

$$\Longrightarrow \quad \pi_0 A_\infty \leq X_0 \quad \text{so} \quad A_\infty < \infty \quad \text{a.e.}$$

More generally $\pi_n A_\infty \leq M_n$ for all n. So we get the decomposition

$$X_n = \underbrace{(M_n - \pi_n A_\infty)}_{\text{pos. mart.}} + \underbrace{(\pi_n A_\infty - A_n)}_{\text{pos. supermart}}$$

The positive supermart. $\bar{X}_n = \pi_n A_\infty - A_n$ is called the potential of the increasing predictable process A_n. Such potentials are characterized by $\pi_0 \bar{X}_n \downarrow 0$ a.e. as $n \to \infty$. The above is called the Riesz decomposition.

Ex. Given B_n an increasing sequence of positive r.v. (not nec. adapted) such that $\pi_0 B_\infty < \infty$ a.e. put

$$\bar{X}_n = \pi_n (B_\infty - B_n)$$

This is a potential.

Ex: Take Markov chain $\Omega = E^{\mathbb{N}}$, E = state space, defined by transition probability $p(x,\cdot)$ for each $x \in E$. I am thinking of E as discrete. On Ω goes the probability P_x such that on Ω_n it is

$$p(x_0, \ldots, x_n) = \delta_x(x_0)\, p(x_0, x_1) \cdots p(x_{n-1}, x_n).$$

Given $f : E \to \mathbb{R}$ we get a process

$$f_n = f(X_n) \qquad X_n : \Omega \to E \quad n\text{th proj.}$$

and one has

$$\pi_{n-1} f_n = (Pf)_{n-1} \qquad (Pf)(x) = \int_y p(x,y) f(y)$$

If f is superharmonic: $Pf \le f$, then f_n is a supermart. Assume $f \ge 0$ and ~~lets~~ compute the Doob decomp.

$$f_n = M_n - A_n$$

$$A_n - A_{n-1} = f_{n-1} - \pi_{n-1} f_n = (f - Pf)_{n-1}$$

$$A_\infty - A_n = (f - Pf)_n + (f - Pf)_{n+1} + \cdots$$

$$\tilde{X}_n = \pi_n (A_\infty - A_n) = (f - Pf)_n + (P(f - Pf))_n + (P^2(f - Pf))_n + \cdots$$

$$= \left(\left(\sum_{n \ge 0} P^n \right)(1 - P) f \right)_n$$

We have to be a bit careful with the limits.

$$f \ge Pf \quad + \quad P \text{ positive matrix} \implies$$

$$f \ge Pf \ge P^2 f \ge \cdots \qquad \text{so} \quad \lim_{n \to \infty} P^n f \quad \text{exists}$$

and it is harmonic. Then

$$\sum_{k=0}^{n-1} P^k (f - Pf) = f - P^n f \quad \uparrow \quad f - P^\infty f$$

Note that the Riesz decomposition amounts to

$$f = \underbrace{(P^\infty f)}_{\text{pos harm.}} + \underbrace{(f - P^\infty f)}_{\text{pos. superharm + killed by } \lim P^n}$$

<u>Dubin's inequality</u>. 1) f_n positive supermartingale, $a>0$
$\Longrightarrow$

$$\pi_0 \, \mathbf{1}_{\{\sup f_n \geq a\}} \leq \min\left(\tfrac{1}{a} f_0, 1\right)$$

2) f_n pos. supermart, $0<a<b<\infty$

$$\pi_0 \, \mathbf{1}_{\{\exists \geq k \text{ upcrossings of } (a,b)\}} \leq \left(\frac{a}{b}\right)^k \min\left(\tfrac{1}{a} f_0, 1\right).$$

Why this is true: Think in terms of discrete prob. spaces, and suppose we have a tower

$$\Omega_0 \twoheadleftarrow \Omega_1 \twoheadleftarrow \Omega_2 \twoheadleftarrow \cdots\cdots \leftarrow \Omega$$

with f_n defined on Ω_n. When we compute π_0 we get a fn. on Ω_0 which gives the relative or conditional expectation. The first inequality says that given $x_0 \in \Omega_0$, the probability of $\sup f_n \geq a$ over events x beginning with x_0 is $\leq \frac{1}{a} f_0(x_0)$ (and ≤ 1 of course).

Now suppose we want the probabity of ≥ 1 upcrossings, again over events x starting with x_0. So introduce stopping times

$$\nu_1(x) = \inf\{n \mid f_n(x) \leq a\}$$
$$\nu_2(x) = \text{"} \ \{n \mid n > \nu_1(x),\ f_n(x) \geq b\}$$

I have to think in terms of Ω consisting of (x_n), $x_n \in \Omega_n$, i.e. $\Omega = \varprojlim \Omega_n$. I want to calculate the probability that $\nu_2 < \infty$. Actually it seems to be possible to define $\Omega_{\nu_2<\infty}$ as $\coprod_{n<\infty}$ Image of $\nu_2 = n$ in Ω_n, and similarly Ω_{ν_1}. Then one has a map $\Omega_{\nu_2<\infty} \to \Omega_{\nu_1<\infty}$ and we want to compute the probability of $\nu_2<\infty$ by bounding the relative probability of the fibres. Thus given $x_0, \dots, x_k$ with $f_k(x_k) \leq a$, the conditional probability that the sequence gets above b later on is bounded by (from 1))

$$\min\left(\frac{1}{b} f_k(x_k), 1\right) \leq \frac{a}{b}$$

This certainly will handle all the future upcrossings ~~at all times we will be applying these~~ however at the first upcrossing we can argue as follows. On the set where $f_0(x_0) \leq a$ we use the bound $\min\left(\frac{1}{b} f_0(x_0), 1\right)$ and on the set where $f_0(x_0) > a$ we just use $\frac{a}{b}$. $= \frac{f_0(x_0)}{b}$. Thus the relative prob. is

$$\begin{cases} \dfrac{f_0(x_0)}{b} & f_0(x_0) \leq a \\[2ex] \dfrac{a}{b} & f_0(x_0) > a \end{cases} \quad = \quad \frac{a}{b} \min\left(\frac{1}{a} f_0(x_0), 1\right)$$

April 17, 1985

Let's review the basics concerning upcrossing inequalities for martingales.

One starts with an increasing family $\mathcal{B}_0 \subset \mathcal{B}_1 \subset \cdots$ of σ-fields and an **adapted** process $\{X_n, n \geq 0\}$ which is a sequence of r.v. with X_n meas rel to $\mathcal{B}_n$. A basic idea is that given a stopping time ν one can define a σ-field $\mathcal{B}_\nu$ and a r.v. X_ν. Thereby one extends notions from integers to stopping times.

(To be more precise about the value ∞. Suppose $\mathcal{B}_\infty$ given containing all $\mathcal{B}_n$. Then

$$\mathcal{B}_\nu = \{ A \mid A \cap \{\nu \leq n\} \in \mathcal{B}_n \quad \text{for all } n \}$$

is defined without problem. However $X_\nu(x) = X_{\nu(x)}(x)$ is defined only if $\nu(x) < \infty$ a.e., or if an X_∞ is given, or if X_n has a limit X_∞.)

Then the definition of martingale:

$$\Pi_n X_m = X_n \qquad n < m$$

can be extended to stopping times

$$\Pi_S(X_T) = X_S \qquad \text{if } S \leq T.$$

and similarly for a submartingale we have

$$\Pi_S(X_T) \geq X_S.$$

Doob's upcrossing inequality: Let X be a submartingale, let U_a^b be the r.v. giving the no. of upcrossings of the interval (a,b) by X. Then

$$\int U_a^b \leq \frac{1}{b-a} \lim_n \int (X_n - a)^+$$

<u>Proof:</u> One can suppose $a = 0$, and one can replace X_n by X_n^+ ($x \mapsto x^+$ is convex and a convex fn. preserves submartingales (modulo int. hypothesis) by Jensen) So one can suppose X_n is a pos. submart.

Think of a submartingale as a favorable game. Consider the following betting strategy.

Better think of X_n as the value of a stock at time n. Here's the strategy. If X_n falls below a we place an order to buy 1 share which we then hold on to until the value rises above b. This isn't quite right.

Maybe I should think of $X_n - X_{n-1}$ as the result of the n-th play. Each time we have to decide how much to bet. If I want to think in terms of stocks I buy at the beginning and sell at the end of each day. I then have to decide whether to do this process for the coming day.

So the strategy is to begin buying + selling, when X_n first drops below a, and to stop when it first rises above b. This gives me a total amount Z_n of $\geq$ the number of upcrossings $\times b$. Here use $X_n \geq 0$, $a = 0$

$$\therefore \; b \int (U_a^b)_n \leq \int Z_n$$

But now because the game is favorable, the more often one plays the better the expectation. This means that

$$\int Z_n \leq \int X_n$$

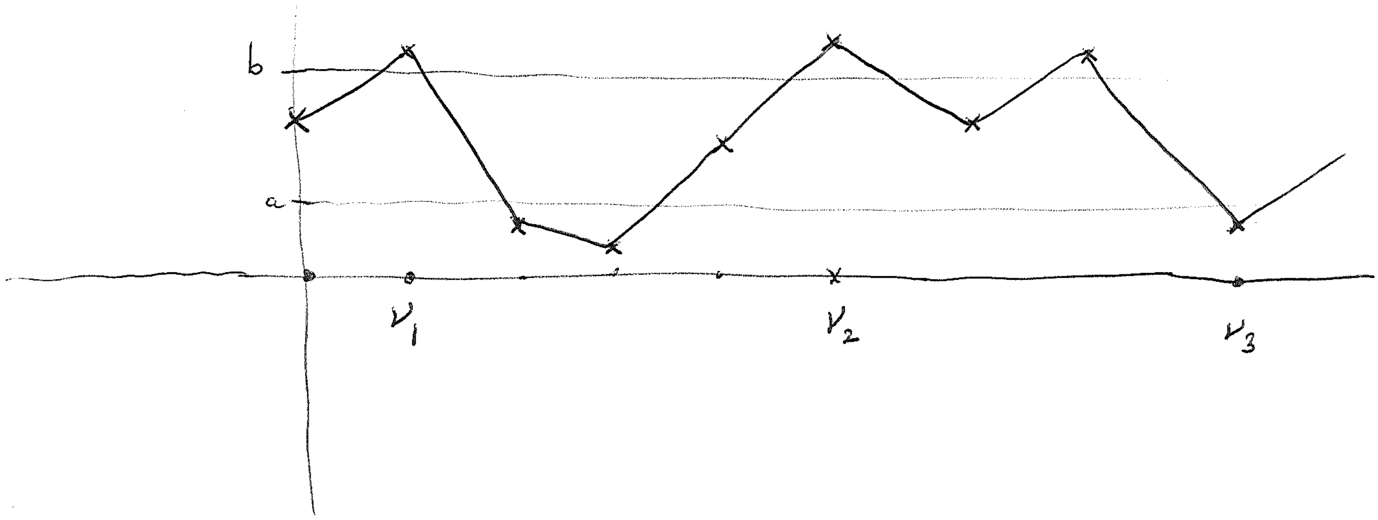
Q.E.D. by letting $n \longrightarrow +\infty$.

Doob's upcrossing inequality for submartingales (assumed integrable, i.e. $X_n \in L^1$ for each $n \geq 0$, so that the conditional expectation π_n has a meaning.)

Given an open interval (a,b) define functions $\nu_1, \nu_2, \ldots$ from Ω to $\mathbb{N} \cup \{+\infty\}$ by

$$\nu_1 = \inf \{ n \geq 0 \mid X_n \leq a \}$$
$$\nu_2 = \inf \{ n \geq \nu_1 \mid X_n \geq b \}$$
$$\nu_3 = \inf \{ n \geq \nu_2 \mid X_n \leq a \} \qquad \text{etc.}$$

Picture



Then the ν_k are stopping times, i.e. $\{\nu_k \leq n\} \in \mathcal{B}_n$ (in words: to determine whether $\nu_k \leq n$ one needs only look at the values of $X_0, \ldots, X_n$)

Now put

$$\beta_n = \text{largest } k \text{ such that } \nu_{2k} \leq n.$$

Then β_n is the number of upcrossings of (a,b) in time $\leq n$.

The result is then

<u>Doob's upcrossing inequality for submartingales:</u>

$$\int \beta_n \leq \frac{1}{b-a} \int (X_n-a)^+ - (X_0-a)^+$$

I want next to describe Doob's proof which is based on the idea that a submartingale is a favorable game and betting. First note that by replacing X_n by $X_n - a$ we can suppose $a=0$. Also because $x \mapsto x^+$ is convex and Jensen's inequality X_n^+ is also a submartingale. (Or more elementary is:

$$\underbrace{\pi_{n-1}(X_n^+)}_{\geq 0} - \underbrace{\pi_{n-1}(X_n^-)}_{\geq 0} = \pi_{n-1}(X_n) \geq X_{n-1} \implies X_{n-1}^+ \leq \pi_{n-1}(X_n^+)).$$

Replacing X_n by X_n^+ doesn't change the functions ν_i, so we can suppose $X_n \geq 0$.

Now consider the following betting [redacted] process. We think of the process X_n [redacted] mainly in terms of its increments $X_n - X_{n-1}$; thus $X_n - X_{n-1}$ represents the change in going from time $n-1$ to time n. A betting scheme consists of giving random [redacted] functions V_n, $n \geq 0$. V_n is the amount one bets on the increment $X_n - X_{n-1}$, and one assumes V_n is measurable relative to $\mathcal{B}_{n-1}$, i.e. the [redacted] amount bet on $X_n - X_{n-1}$ depends only upon one's knowledge of the process up through time $n-1$. Such a process is called <u>predictable</u>. (By convention it is useful to allow V_0 to be a constant).

The result of the betting is the process

$$(V \cdot X)_n = V_0 X_0 + V_1(X_1 - X_0) + \cdots + V_n(X_n - X_{n-1})$$

This gives the net gain or loss at time n.

Example: Let $\nu : \Omega \longrightarrow \mathbb{N} \cup \{+\infty\}$ be a stopping time. and take

$$V_n(\omega) = \begin{cases} 1 & n \leq \nu(\omega) \\ 0 & n > \nu(\omega) \end{cases}$$

or

$$V_n = 1_{n \leq \nu}. \quad \text{Note that}$$

$$V_n = 1 - 1_{\nu < n} = 1 - 1_{\nu \leq n-1}$$

is measurable relative to $\mathcal{B}_{n-1}$. Thus $\{V_n\}$ is predictable. Clearly

$$(V \cdot X)_n = X_0 + (X_1 - X_0) + \cdots + (X_{\nu \wedge n} - X_{\nu \wedge n - 1})$$

$$= X_{\nu \wedge n}.$$

This process represents betting 1 until the stopping time is reached and then stopping.

Next we want to derive the fact if we are going to bet amounts V_n with $0 \leq V_n \leq 1$ on a submartingale (which is supposed to be a favorable game), then our expectation is greatest when we bet 1 all the time. Put $Y_n = (V \cdot X)_n$ whence

$$Y_n - Y_{n-1} = V_n (X_n - X_{n-1}).$$

Then

$$\pi_{n-1} Y_n - Y_{n-1} = \pi_{n-1}\{V_n (X_n - X_{n-1})\}$$

$$= \underbrace{V_n}_{0 \leq \ \leq 1} \cdot \underbrace{(\pi_{n-1}(X_n) - X_{n-1})}_{\geq 0} \leq \pi_{n-1} X_n - X_{n-1}$$

so

$$\int Y_n - \int Y_{n-1} \leq \int X_n - \int X_{n-1}$$

or

$$\int Y_n - Y_0 \leq \int X_n - X_0$$

Now apply this to the ~~[illegible]~~ following betting process in the case of a positive submartingale.
We bet nothing until the time

$$\nu_1 = \inf\{n \mid X_n = 0\}$$

then we bet 1 each time until the time

$$\nu_2 = \inf\{n > \nu_1 \mid X_n \geq b\}$$

then zero until the time ν_3, etc. Thus

$$V_n = \begin{cases} 1 & \text{if} \quad \nu_{2k-1} < n \leq \nu_{2k} \quad \text{for some } k \\ 0 & \text{otherwise} \end{cases}$$

$$V_n = \left(1_{\nu_2 \leq n} - 1_{\nu_1 \leq n}\right) + \left(1_{\nu_4 \leq n} - 1_{\nu_3 \leq n}\right) + \cdots$$

$$Y_n = \left(X_{\nu_2 \wedge n} - X_{\nu_1 \wedge n}\right) + \left(X_{\nu_4 \wedge n} - X_{\nu_3 \wedge n}\right) + \cdots$$

$$= \begin{cases} 0 & n \leq \nu_1 \\ X_n - X_{\nu_1} & \nu_1 < n \leq \nu_2 \\ X_{\nu_2} - X_{\nu_1} & \nu_2 < n \leq \nu_3 \\ X_n - X_{\nu_3} + (X_{\nu_2} - X_{\nu_1}) & \nu_3 < n \leq \nu_4 \quad \ldots \end{cases}$$

Note $X_{\nu_{odd}} = 0$, $X_{\nu_{ev}} \geq b$ (when these ν's are finite)

Thus $Y_n \geq b\beta_n$ for all n. (In words, we wait to the process drops to 0 then we bet 1 until it rises $\geq b$ and stop, then start when it next falls to zero, etc. Our gain is $\geq b \cdot$ no. of upcrossings)

Putting things together we get

$$b\int \beta_n \leq \int Y_n \leq \int X_n - X_0$$

which is exactly Doob's upcrossing inequality.

Next I want to describe Neveu's proof which is based on the following switching lemma.

<u>Lemma:</u> Let X_n and Y_n be submartingales and let ν be a stopping time. Assume $X_\nu \leq Y_\nu$ on $\nu < \infty$. Then

$$Z_n = X_n 1_{n<\nu} + Y_n 1_{n \geq \nu}$$

is a submartingale.

Proof. $$Z_n = X_n \underbrace{1_{n \leq \nu}}_{1 - 1_{n>\nu} \in \mathcal{B}_{n-1}} + \underbrace{(Y_n - X_n) 1_{n=\nu}}_{\geq 0} + Y_n \underbrace{1_{n>\nu}}_{\in \mathcal{B}_{n-1}}$$

so $$\pi_{n-1} Z_n = \underbrace{\left(\pi_{n-1} X_n\right)}_{\geq X_{n-1}} 1_{n \leq \nu} + \underbrace{\pi_{n-1}\left\{(Y_n - X_n) 1_{n=\nu}\right\}}_{\geq 0} + \underbrace{\pi_{n-1}(Y_n)}_{\geq Y_{n-1}} 1_{n>\nu}$$

$$\geq X_{n-1} 1_{n \leq \nu} + Y_{n-1} \cdot 1_{n>\nu}$$

$$= X_{n-1} 1_{n-1<\nu} + Y_{n-1} 1_{n-1 \geq \nu} = Z_{n-1}$$

Now define the stopping times as before on p. 448:

$$\nu_1 = \inf\{n \geq 0 \mid X_n \leq a\}$$
$$\nu_2 = \inf\{n \geq \nu_1 \mid X_n \geq b\}$$
$$\cdots$$

and define Z_n as follows

$$Z_n = \begin{cases} X_n - a & 0 \leq n < \nu_1 \\ 0 & \nu_1 \leq n < \nu_2 \\ (X_n - a) - (b-a) & \nu_2 \leq n < \nu_3 \end{cases}$$

$$= \begin{cases} -(b-a) & \nu_3 \le n < \nu_4 \\ (X_n - a) - 2(b-a) & \nu_4 \le n < \nu_5 \end{cases}$$

in general

$$Z_n = \begin{cases} (X_n - a) - (b-a)\beta_n & \text{on } \bigcup_{k\geq 0} \{\nu_{2k} \le n < \nu_{2k+1}\} \\ -(b-a)\beta_n & \bigcup_{k \geq 0} \{\nu_{2k+1} \le n < \nu_{2k+2}\} \end{cases}$$

The switching lemma implies that Z_n is a submartingale. Note

$$Z_0 = (X_0 - a)^+$$
$$Z_n \le (X_n - a)^+ - (b-a)\beta_n$$

and so we have

$$\int Z_n \geq \int Z_0$$

or

$$(b-a)\int \beta_n \le \int (X_n - a)^+ - (X_0 - a)^+$$

Dubin's inequality for positive supermartingales (not necessarily integrable). Let $0 < a < b < \infty$. Let β be the number of upcrossings of (a, b). Then

$$\Pi_0\left(1_{\{\beta \geq k\}}\right) \le \left(\frac{a}{b}\right)^k \min\left(\frac{X_0}{a}, 1\right)$$

Proof after Neveu: Introduce stopping times

$$\nu_1 = \inf\{n \geq 0 \mid X_n \le a\}$$
$$\nu_2 \qquad \{n \geq \nu_1 \mid X_n \geq b\}$$

etc. down thru ν_{2k}, and then using switching to see that

$$Z_n = \begin{cases} 1 & 0 \le n < \nu_1 \\ \dfrac{X_n}{a} & \nu_1 \le n < \nu_2 \\ \dfrac{b}{a} & \nu_2 \le n < \nu_3 \\ \dfrac{b}{a}\,\dfrac{X_n}{a} & \nu_3 \le n < \nu_4 \\ \cdots\cdots\cdots & \\ \left(\dfrac{b}{a}\right)^k & \nu_{2k} \le n \end{cases}$$

is a positive supermartingale. So

$$\min\left(\frac{X_0}{a}, 1\right) = Z_0 \geq \Pi_0 Z_n \geq \Pi_0\left(Z_n \cdot 1_{\beta \geq k}\right) = \left(\frac{b}{a}\right)^k \Pi_0\left(1_{\beta \geq k}\right)$$

Next consider the maximal inequality for positive supermartingales

$$\Pi_0\, 1\left\{\sup_{k \le n}(X_k) \geq a\right\} \leq \min\left(\frac{X_0}{a}, 1\right)$$

By a limiting argument it is enough to show

$$\Pi_0\, 1_{\left\{\sup_{k\le n} X_k > a\right\}} \leq \frac{X_0}{a}$$

I want to do this the way I understood earlier and then by Neveu's switching method. Let

$$\nu = \inf\{n \geq 0 \mid X_n > a\}$$

so that $\left\{\sup_{k \le n} X_k > a\right\} = \{\nu \le n\}$. Then we argue as follows

$$X_0 \geq \Pi_0(X_{\nu \wedge n}) \geq \Pi_0(X_{\nu \wedge n} \cdot 1_{\nu \leq n}) \geq a\, \Pi_0(1_{\nu \leq n}).$$

The middle $\geq$ results from the positivity of $X_{\nu \wedge n}$, the last inequality from the fact that $X_\nu \geq a$ on $\nu \leq n$. The key step then is the first inequality where the fact that X_n is a supermartingale is used. It suffices to show that $\Pi_{n-1}(X_{\nu \wedge n}) \leq X_{\nu \wedge (n-1)}$ for any stopping time ν, i.e. that $X_{\nu \wedge n}$ is a supermartingale. But this follows from the fact that $X_{\nu \wedge n} = (V \cdot X)_n$ where $V_n = 1_{n \leq \nu}$ and that such a transform for $V \geq 0$ is always a supermartingale.

Specifically in the case of interest

$$X_{\nu \wedge n} - X_{\nu \wedge (n-1)} = 1_{n \leq \nu} \cdot (X_n - X_{n-1})$$

so

$$\Pi_{n-1} X_{\nu \wedge n} - X_{\nu \wedge (n-1)} = \underbrace{1_{n \leq \nu}}_{\geq 0} (\underbrace{\Pi_{n-1} X_n - X_{n-1}}_{\leq 0}) \leq 0.$$

Next consider Neveu's proof. With ν as before set

$$Z_n = \begin{cases} X_n & n < \nu \\ a & n \geq \nu. \end{cases}$$

By switching this is a supermartingale so

$$\underset{\min(X_0, a)}{\overset{\parallel}{Z_0}} \geq \Pi_0(Z_n) \geq \Pi_0(Z_n \cdot 1_{\nu \leq n}) = a\, \Pi_0(1_{\nu \leq n})$$

so $\Pi_0(1_{\nu \leq n}) \leq \min\left(\frac{X_0}{a}, 1\right)$ ∎

The upcrossing inequality for positive supermartingales leads easily to the fact they converge a.e. Set $X_\infty = \lim X_n$: Fatou's lemma says

$$\pi_n(X_\infty) = \pi_n\left(\underline{\lim}\, X_p\right) \leq \underline{\lim}\, \pi_n(X_p) \leq X_n$$

and so $X_\infty < \infty$ a.e. outside the set where all X_n are ∞. So some $X_n \in L^1 \implies X_\infty \in L^1$.

Now let us turn to the submartingale convergence thm:

<u>Thm</u>: Assume X_n is an integrable submartingale such that $\sup \int X_n^+ < \infty$. Then X_n converges a.e. to an integrable X_∞.

Proof. X_n^+ is a positive submartingale, so $\pi_n(X_p^+)$ is increasing for $p \geq n$ and we can define

$$M_n = \limsup_p \pi_n(X_p^+) \qquad (= \lim \text{ except where it is } \infty.)$$

By monotone convergence

$$\int M_n = \lim \int X_p^+ < \infty$$

so each $M_n \in L^1$. Also

$$\pi_{n-1}(M_n) = \lim_p \pi_{n-1}\pi_n X_p^+ = \lim_p \pi_{n-1} X_p^+ = M_{n-1}$$

so M_n is a positive martingale.

By the convergence theorem for pos. supermartingales

$$M_n \longrightarrow M_\infty \quad \text{a.e.}$$

where

$$\int M_\infty \leq \int M_n = \lim \int X_p^+, \text{ so } M_\infty \in L^1.$$

Next we have

$$M_n = \lim \pi_n(X_p^+) \geq X_n^+ \geq X_n$$

so if we put $Y_n = M_n - X_n$, then $Y_n \in L^1$ and $Y_n \geq 0$. Clearly because M_n is a mart. + X_n is a submart, Y_n is a supermart. Thus by the supermart. conv. thm. $Y_n \to Y_\infty$ a.e. where $Y_\infty \in L^1$. Thus $X_n = M_n - Y_n \longrightarrow M_\infty - Y_\infty$ a.e. and this belongs to L^1. QED.

In general

$$\int X = \int X^+ - \int X^-$$

$$\int |X| = \int X^+ + \int X^-$$

$$2\int X^+ = \int |X| + \int X$$

For a martingale $\int X_n$ is constant + finite so

$$\sup \int X_n^+ < \infty \iff \sup \int |X_n| < \infty .$$

Thus a martingale bounded in L^1 converges a.e. to an element of L^1.

Doob proves: If X_n is a martingale bdd in L^p, then X_n converges in L^p for $p > 1$. For $p=1$ false, but true if X_n is a <u>regular</u> integrable martingale:

$$\sup_N \int_{|X_N| > a} |X_n| \downarrow 0 \qquad \text{as} \quad a \uparrow \infty$$

(satisfied if $\sup |X_n| \in L^1$ and this follows from $\sup_N \int |X_n| \log^+ |X_n| < \infty$)

April 20, 1985

What is a martingale? To simplify we consider the discrete case where the set of times in $\mathbb{N}$. A martingale is a stochastic process $\{X_n, n \in \mathbb{N}\}$. This means we are given a probability space $(\Omega, \mathcal{B}, P)$ on which the X_n are random variables. Further we are given a filtration $\mathcal{B}_0 \subset \mathcal{B}_1 \subset \cdots$ of the σ-algebra $\mathcal{B}$, where $\mathcal{B}_n$ is the algebra of events up to time n. I like to think of Ω as the inverse limit of a tower

$$\Omega_0 \twoheadleftarrow \Omega_1 \twoheadleftarrow \Omega_2 \twoheadleftarrow \cdots\cdots$$

of probability spaces and of $\mathcal{B}_n$ as the measurable subsets in Ω_n. A point of Ω_n is a possible history up to time n.

With this notation an integrable martingale is sequence of integrable real r.v.'s X_n which is adapted to the filtration $\{\mathcal{B}_n\}$ (i.e. X_n is measurable relative to $\mathcal{B}_n$), such that the conditional expectation of X_n relative to $\mathcal{B}_m$ where $m<n$ is X_m. Equivalently, the conditional expectation relative to $\mathcal{B}_{n-1}$ of the n-th increment $X_n - X_{n-1}$ is zero. Thus if we know the history up to time $n-1$ the average value of $X_n - X_{n-1}$ is zero.

An example of an integrable martingale is

$$X_n = Y_0 + \cdots + Y_n$$

where the Y_j are independent integrable r.v.'s of mean zero.

Here are the martingale convergence theorems.

1) A bounded L^1-martingale X_n (means $\sup_N \int |X_n| < \infty$) converges a.e. The limit X_∞ is in L^1.

Recall $L^p \subset L^1$ for $1 \leq p$, so that a bounded L^p martingale X_n has an a.e. limit X_∞ by 1).

2) <u>If</u> $1 < p < \infty$, [redacted] <u>and</u> X_n <u>is a bounded</u> L^p <u>martingale</u>, <u>then</u> $X_n \longrightarrow X_\infty$ <u>in</u> L^p.

<u>Remarks:</u>

a) For $p = 2$ it is very easy to see [redacted] that an L^2 bounded martingale converges in L^2, because the conditional expectation operator coincides with the orthogonal projection: $L^2(\Omega) \longrightarrow L^2(\Omega_n)$. Thus for an L^2-martingale the increments $X_n - X_{n-1}$ are mutually orthogonal and

$$\|X_n\|_2^2 = [redacted]\ \|X_0\|_2^2 + \sum_{k=1}^{n} \|X_k - X_{k-1}\|_2^2$$

and so if this is bounded the series

$$X_\infty = X_0 + \sum_{1}^{\infty} (X_k - X_{k-1})$$

converges in L^2.

b) Since $x \longmapsto |x|^p$ is convex for $p \geq 1$, it follows from Jensen that $|X_n|^p$ is a submartingale, hence $\int |X_n|^p$ (also $\|X_n\|_p$) is an increasing sequence

The proof of 2) is based on <u>Doob's "maximal" inequality</u>: If $\{X_n\}$ is a positive (integrable) submartingale, and $X_n^* = \sup_{k \leq n} X_k$, then

$$P\{X_N^* \geq a\} \leq \frac{1}{a} \int_{\{X_N^* \geq a\}} X_N \, dP \qquad a > 0.$$

and hence

$$P\{X_\infty^* \geq a\} \leq \frac{1}{a} \sup_N \int X_n \qquad X_\infty^* = \sup_N X_n$$

The reason I ~~[illegible]~~ record this is that a key idea is whether X_∞^* is <u>integrable</u> ~~[illegible]~~

Let X_n be an L^p bounded martingale, $1 \le p < \infty$. From the maximal inequality

$$P\{\sup_{k \le n} |X_k| > a\} \le \frac{1}{a} \int_{\{\sup_{k \le n} |X_k| > a\}} |X_n|$$

(or $\ge$)

one can deduce that $\sup |X_n| \in L^p$ for $p > 1$. (In fact

$$\circledast \qquad \|\sup_{k \le n} |X_k|\|_p \le \frac{p}{p-1} \|X_n\|_p \;)$$

By the martingale convergence thm. one has $X_n \to X_\infty$ a.e. By dominated convergence $|X_n| \le \sup_N |X_n| \in L^p$, we see that $X_n \to X_\infty$ in L_p.

Thus even when $p = 1$ we have

$$\sup |X_n| \in L^1 \implies X_n \to X_\infty \text{ in } L^1$$

~~The actual behavior at p=1 is very delicate~~

Variant of $\circledast$

$$\int \sup_{k \le n} |X_k| \le \frac{e}{e-1} \left(1 + \int |X_n| \log^+ |X_n|\right)$$

which shows that $\int |X_n| \log^+ |X_n|$ bounded $\implies X_n \to X_\infty$ in L^1.

The actual behavior at $p = 1$ is very delicate and there is an extensive development analogous to Fefferman's theorem that H^1 and BMO are dual. Also Littlewood-Paley theory is involved.

I make an attempt to summary the theory. First we must describe the space $\mathcal{M}^p$ (or $\mathcal{H}^p$). For $p>1$ it consists of bounded L^p martingales with either of the equivalent norms

$$\sup \|X_n\|_p \qquad \text{or} \qquad \|\sup |X_n|\,\|_p$$

Thus $\mathcal{M}^p$ should be the same as $L^p(\Omega)$ (assuming $\bigcup L^p(\Omega_n)$ is dense in $L^p(\Omega)$.

Now we take up $\mathcal{H}^1$. This consists of L^1-bounded martingales X (with $X_0=0$ to simplify) with either of two equivalent norms. (The equivalence is non-trivial.) The first is the "maximal" norm:

$$\int \sup_N |X_n|$$

and the second is the "quadratic" norm

$$\int \sqrt{V} = \int \Big(\sum_N \pi_n\big((X_{n+1}-X_n)^2\big)\Big)^{1/2}$$

(Remark: Littlewood - Paley theory concerns Fourier series being summed in a "dyadic" way. There is an analogue for martingales in which the interesting L^p-gadget is

$$\int\Big[\Big(\sum \pi_n (X_n - X_{n-1})^2\Big)^{1/2}\Big]^p = \int (\sqrt{V})^p$$

In any case there is an equivalence of norms

$$\|\sqrt{V}\|_p \quad \text{and} \quad \|\sup |X_n|\,\|_p$$

for all $p \geq 1$.)