MATH 585 – TOPICS IN MATHEMATICAL PHYSICS – FALL 2006

MATHEMATICS OF MEAN FIELD SPIN GLASSES AND THE REPLICA METHOD

LECTURE 6: DE FINETTI'S THEOREM I

S. STARR
MATHEMATICS DEPARTMENT, UNIVERSITY OF ROCHESTER

CONTENTS

1. ABSTRACT NONSENSE

In this lecture, I will freely state some topological and analytical facts. In a companion lecture, I will give the complete proofs of the facts needed in a concrete setting. Most of these facts follow from high-power general theorems that may use, for example, Zorn's lemma, the Deus-ex-machina of choice in point-set topology. Of course, in concrete situations, one should never have to use that. If I can figure them out, I will give the concrete proofs instead of referring to general theorems. But also, we will not have time or inclination to discuss that promised lecture in class. So basically, just take these facts at face-value, only one of them contains anything important.

A good reference for point-set topology and basic analysis, as well as measure theory and probability, is Dudley's monograph, [3]. Other books on analysis are Royden [6], Kelley [5], Choquet [1], and Dunford and Schwartz [2], and Simon [7].

Let $\Omega = \{+1, -1\}$. The only thing that is important for us about this set is that it can be considered as a compact metric space. Namely, let the metric be $d_1(\sigma, \sigma') = \frac{1}{2}|\sigma - \sigma'|$. Then this leads to the discrete topology: every one of the four possible subsets of Ω is both open and closed. Of course, every set is also measurable. Let $\mathscr{M}_1(\Omega)$ be the set of all Borel probability measures on Ω.

Let us introduce some notation now. Note that $\mathscr{M}_1(\Omega)$ is isomorphic to $[0,1]$. Namely, given $p \in [0,1]$, define $\alpha_p \in \mathscr{M}_1(\Omega)$ such that

$$\alpha_p\{+1\} = p \quad \text{and} \quad \alpha_p\{-1\} = 1 - p\,.$$

Since we are considering Ω as a compact metric space, we can ask what are the continuous functions into $\mathbb{R}$. They are all functions. In this case $\mathscr{C}(\Omega)$, is isomorphic to $\mathbb{R}^2$. Namely,

Date: October 26, 2006.

a continuous function is given by a pair $(f(+1), f(-1))$, The topology on $\mathscr{C}(\Omega)$ is the sup-norm topology. So $\|f-g\|_\infty = \max\{|f(+1)-g(+1)|, |f(-1)-g(-1)|\}$.

Let $\mathscr{M}_\pm(\Omega)$ denote the set of signed measures. So for $\mu \in \mathscr{M}_\pm(\Omega)$, one has that $\mu\{+1\}$ and $\mu\{-1\}$ are arbitrary numbers in $\mathbb{R}$. Thus $\mathscr{M}_\pm(\Omega)$ is also isomorphic to $\mathbb{R}^2$. Given $\mu \in \mathscr{M}_\pm(\Omega)$ and $f \in \mathscr{C}(\Omega)$, one can define a pairing

$$\langle f, \mu\rangle \,=\, \int_\Omega f\, d\mu \,=\, f(+1)\mu\{+1\} + f(-1)\mu\{-1\}\,.$$

As vectors, this is the inner-product. One puts the weak topology on $\mathscr{M}_\pm(\Omega)$. This means that a sequence $\mu_1, \mu_2, \dots \in \mathscr{M}_\pm(\Omega)$ converges to a measure $\mu \in \mathscr{M}_\pm(\Omega)$ if and only if, for each $f \in \mathscr{C}(\Omega)$, one has

$$\lim_{n\to\infty} \int_\Omega f\, d\mu_n \,=\, \int_\Omega f\, d\mu\,.$$

Of course, on a finite-dimensional vector space like $\mathbb{R}^2$, essentially all nondegnerate topologies are equivalent. But this is not true for infinite-dimensional vector spaces.

For each $N \in \mathbb{Z}_{>0}$, let Ω^N be the usual product space. Also, define a metric $d_N(\cdot,\cdot)$ on Ω^N by

$$d_N(\sigma, \sigma') \,=\, \sum_{n=1}^N 2^{-n}|\sigma_n - \sigma'_n|\,,$$

for $\sigma = (\sigma_1, \dots, \sigma_N)$ and $\sigma' = (\sigma'_1, \dots, \sigma'_N)$ in Ω^N. Then, once again, the topology associated to this metric is discrete. All sets are open and closed. Let $\mathscr{M}_1(\Omega^N)$, $\mathscr{C}(\Omega^N)$ and $\mathscr{M}_\pm(\Omega^N)$ be the sets of all Borel probability measures on Ω^N, all continuous functions from Ω^N to $\mathbb{R}$, and all signed Borel measures on Ω^N. Of course, in this case, because the topology is discrete, all measures are Borel, and all functions are continuous. Given $p \in [0,1]$, define $\alpha_p^{\otimes N} \in \mathscr{M}_1(\Omega^N)$ to be the product measure. In other words,

$$\alpha_p^{\otimes N}\{\sigma\} \,=\, \prod_{n=1}^N \alpha_p\{\sigma_n\}\,.$$

There is, of course, a better way of saying this. If $\mathsf{X} = (\mathsf{X}_1, \dots, \mathsf{X}_N) \in \Omega^N$ is random and distributed according to $\alpha_p^{\otimes N}$, then this is exactly the same as that $\mathsf{X}_1, \dots, \mathsf{X}_N$ are i.i.d. and distributed by α_p. So we will say $\alpha_p^{\otimes N}$ is an i.i.d. product measure.

Let Ω^∞ be the set of all infinite sequences $\sigma = (\sigma_1, \sigma_2, \dots)$ with all $\sigma_1, \sigma_2, \dots \in \Omega$. Let Ω^∞ have the product topology. This means the following. For each $N \in \mathbb{Z}_{>0}$, let $\phi_N : \Omega^\infty \to \Omega^N$ be the standard projection, $\phi_N(\sigma) = (\sigma_1, \dots, \sigma_N)$ for $\sigma \in \Omega^\infty$. Then the product topology on Ω^∞ is the smallest/weakest/coarsest topology such that each of the maps ϕ_N is continuous. In other words, a sub-base of the topology is given by the set of all sets: for $N \in \mathbb{Z}_{>0}$ and $(\sigma'_1, \dots, \sigma'_N) \in \Omega^N$, the set $\phi_N^{-1}(\{(\sigma'_1, \dots, \sigma'_N)\}) = \{\sigma \in \Omega^\infty : \sigma_n = \sigma'_n \text{ for } n = 1, \dots, N\}$. In the context of measure theory, we might be comfortable calling these "cylinder sets". Recall that the definition of sub-base for a topology means that a base for the topology is given by all finite intersections of sets in the sub-base. The definition of base for a topology means that all open sets are arbitrary unions of sets in the sub-base.

Fact 1. *With the topology given, Ω^∞ is a compact topological space. More than this it is metrizable, meaning that there is a metric such that the associated metric topology is the* *same as the product topology. Specifically, let*

$$d(\sigma, \sigma') \,=\, \sum_{n=1}^{\infty} 2^{-n} |\sigma_n - \sigma'_n| \,.$$

Then this is a good metric, compatible with the topology on Ω^∞.

We let $\mathscr{C}(\Omega^\infty)$ denote the Banach space of all continuous functions from Ω^∞ to $\mathbb{R}$, with the sup-norm. Let $\mathscr{M}_1(\Omega^\infty)$ denote the set of all Borel probability measures on Ω^∞. The good topology on $\mathscr{M}_1(\Omega^\infty)$ is the weak topology. Let us again define the pairing between $f \in \mathscr{C}(\Omega^\infty)$ and $\mu \in \mathscr{M}_1(\Omega^\infty)$,

$$\langle f, \mu \rangle \,=\, \int_{\Omega^\infty} f \, d\mu \,.$$

The *weak topology* on $\mathscr{M}_1(\Omega^\infty)$ is the smallest/weakest/coarsest topology on $\mathscr{M}_1(\Omega^\infty)$ such that the map $\langle f, \cdot \rangle : \mathscr{M}_1(\Omega^\infty) \to \mathbb{R}$ is continuous for each $f \in \mathscr{C}(\Omega^\infty)$. This has nice features.

Fact 2. $\mathscr{M}_1(\Omega^\infty) \subset \mathscr{M}_\pm(\Omega^\infty)$*, with the weak topology, is compact and metrizable.*

Note that $\mathscr{M}_1(\Omega^\infty)$ has a convex structure: if μ_1 and μ_2 are both probability measures, and if $\theta \in [0,1]$, then of course $\theta \cdot \mu_1 + (1-\theta)\mu_2$ is also a probability measure. Recall the definition of an extreme point of a convex set C. It is a point $x \in C$ such that for any $\theta \in (0,1)$, the only $y, z \in C$ solving the identity $x = \theta \cdot y + (1-\theta) z$ are $y = z = x$. Also, given a set A in a convex topological space, recall the definition of the closed convex hull:

$$\begin{aligned}\mathrm{cch}(A) \,:=\, \mathrm{cl}(\{x \,:\, &\exists N \in \mathbb{Z}_{>0}\,,\ \exists x_1, \dots, x_N \in A\,,\ \exists \theta_1, \dots, \theta_N \in [0,1] \\ &\text{such that } \theta_1 + \dots + \theta_N = 1 \text{ and } x = \theta_1 x_1 + \dots + \theta_N x_N\})\,,\end{aligned}$$

where $\mathrm{cl}(E) = \overline{E}$ is the closure operation.

Fact 3. *Let K be a compact, convex subset of $\mathscr{M}_1(\Omega^\infty)$. Then K is a subset of the closed convex hull of its extreme points.*

2. De Finetti's theorem

Suppose that $\pi : \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ is an injection. Then we can define a mapping $\pi_* : \Omega^\infty \to \Omega^\infty$ such that

$$\pi_*(\sigma) \,=\, (\sigma_{\pi(1)}, \sigma_{\pi(2)}, \dots)\,.$$

Fact 4. *The mapping $\pi_* : \Omega^\infty \to \Omega^\infty$ is continuous. Therefore, the mapping from $\mathscr{M}_1(\Omega^\infty) \to \mathscr{M}_1(\Omega^\infty)$ given by $\mu \mapsto \mu \circ \pi_*^{-1}$ is (weakly) continuous.*

Definition 2.1 *A measure $\mu \in \mathscr{M}_1(\Omega^\infty)$ is called "exchangeable" if*

$$\mu \,=\, \mu \circ \pi_*^{-1}\,,$$

for every injection $\pi : \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$.

Theorem 2.2 (de Finetti) *Let $\mu \in \mathscr{M}_1(\Omega^\infty)$ be an exchangeable measure. Then there is a unique $\rho \in \mathscr{M}_1([0,1])$ such that*

$$\mu \circ \phi_N^{-1}(\cdot) \;=\; \int_0^1 \alpha_p^{\otimes N}(\cdot)\, d\rho(p)\,,$$

for each $N \in \mathbb{Z}_{>0}$.

3. The Choquet-theoretic proof

The argument of this section is due to Hewitt and Savage [4]. Let K be the set of all exchangeable measures in $\mathscr{M}_1(\Omega^\infty)$. It is obvious that K is convex because the condition $\mu = \mu \circ \pi_*^{-1}$ is linear in μ, for each injection $\pi : \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$.

Lemma 3.1 $\mathscr{E}(K) \subseteq \{\alpha_p^{\otimes\infty} : p \in [0,1]\}$.

Proof. Suppose $\mu \in \mathscr{E}(K)$. Then one of three things happens:

(1) $\mu \circ \phi_1^{-1}\{+1\} = 1$;
(2) $\mu \circ \phi_1^{-1}\{+1\} = 0$; or
(3) $0 < \mu \circ \phi_1^{-1}\{+1\} < 1$.

It is easy to see that (1) implies $\mu = \alpha_1^{\otimes\infty}$, and (2) implies $\mu = \alpha_0^{\otimes\infty}$. Let us prove this, because it is also a warm-up for the argument for case (3).

Suppose μ satisfies (1). Suppose $N \in \mathbb{Z}_{>0}$ and $(\sigma_1, \dots, \sigma_N) \in \Omega^N$ is anything other than $(+1, \dots, +1)$. Then we want to prove that $\mu \circ \phi_N^{-1}\{(\sigma_1, \dots, \sigma_N)\} = 0$. That family of conditions would be sufficient to prove that $\mu \circ \phi_N^{-1}\{(+1, \dots, +1)\} = 1$. Moreover, proving that for each $N \in \mathbb{Z}_{>0}$ is precisely equivalent to: $\mu = \alpha_1^{\otimes\infty}$. Therefore, let k_1 be the first index in $[1, N]$ such that $\sigma_{k_1} = -1$. Then, define the injection $\pi : \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ by $\pi(1) = k_1, \pi(2) = N+1, \pi(3) = N+2, \dots$. By exchangeability and (1), $\mu \circ \pi^{-1} \circ \phi_1^{-1}\{-1\} = \mu \circ \phi_1^{-1}\{-1\} = 0$. But clearly, $\pi^{-1} \circ \phi_1^{-1}\{-1\} \supseteq \phi_N^{-1}\{(\sigma_1, \dots, \sigma_N)\}$, so this proves just what we wanted. A symmetric argument works for case (2).

Suppose μ satisfies (3). Let $\pi : \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ be the injection $\pi(n) = n+1$, for each $n \in \mathbb{Z}_{>0}$. Define two new measures $\mu_1, \mu_2 \in \mathscr{M}_1(\Omega^\infty)$ by

$$\mu_1(\cdot) \;=\; \frac{\mu(\phi_1^{-1}\{+1\} \cap \pi^{-1}(\cdot))}{\mu \circ \phi_1^{-1}\{+1\}} \quad \text{and} \quad \mu_2(\cdot) \;=\; \frac{\mu(\phi_1^{-1}\{-1\} \cap \pi^{-1}(\cdot))}{\mu \circ \phi_1^{-1}\{-1\}}\,,$$

which are well-defined because of (3). Obviously

$$\mu \;=\; \mu \circ \phi_1^{-1}\{+1\} \cdot \mu_1 + \mu \circ \phi_1^{-1}\{-1\} \cdot \mu_2\,,$$

which is a convex combination with $\theta = \mu \circ \phi_1^{-1}\{+1\} \in (0,1)$. Therefore, since μ is supposed to be extremal in K, if it can be proved that $\mu_1, \mu_2 \in K$, it will follow that $\mu_1 = \mu_2 = \mu$. But indeed, it is easy to see that $\mu_1, \mu_2 \in K$. For, if $\pi_1 : \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ is any injection, defining $\pi_2 : \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ so that $\pi_2(1) = 1$ and $\pi_2(n+1) = \pi_1(n)$ for all $n \in \mathbb{Z}_{>0}$, it is easy to see that

$$\phi_1^{-1}\{+1\} \cap \pi^{-1} \circ \pi_1^{-1}(\cdot) \;=\; \pi_2^{-1}\left(\phi_1^{-1}\{+1\} \cap \pi^{-1}(\cdot)\right)\,.$$

Therefore,

$$\begin{aligned}
\mu_1 \circ \pi_1^{-1}(\cdot) &\stackrel{\text{def}}{=} \frac{\mu(\phi_1^{-1}\{+1\} \cap \pi^{-1} \circ \pi_1^{-1}(\cdot))}{\mu \circ \phi_1^{-1}\{+1\}} \\
&= \frac{\mu \circ \pi_2^{-1}(\phi_1^{-1}\{+1\} \cap \pi^{-1}(\cdot))}{\mu \circ \phi_1^{-1}\{+1\}} \\
&= \frac{\mu(\phi_1^{-1}\{+1\} \cap \pi^{-1}(\cdot))}{\mu \circ \phi_1^{-1}\{+1\}} \quad \text{(by exchangeability of } \mu) \\
&= \mu_1(\cdot)\,.
\end{aligned}$$

A symmetric argument works for μ_2 (or observe that by linearity, if $\mu, \mu_1 \in K$, it must be that $\mu_2 \in K$).

Therefore, $\mu_1 = \mu_2 = \mu$. This means that $\mu = \alpha_p \otimes \mu \circ \pi^{-1}$, for $p = \mu \circ \phi_1^{-1}\{+1\}$ and $\pi : \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ the injection above. But, using exchangeability and iterating the argument, this implies that $\mu = \alpha_p^{\otimes\infty}$. □

For each injection, $\pi : \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$, the set $\{\mu \,:\, \mu = \mu \circ \pi_*^{-1}\}$ is closed in $\mathscr{M}_1(\Omega^\infty)$. (in part because $\mu \mapsto \mu \circ \pi_*^{-1}$ is continuous). The intersection of an arbitrary family of closed sets is closed. Therefore, K is closed. So it is a closed subset of the compact set $\mathscr{M}_1(\Omega^\infty)$. Therefore, it is compact. Therefore, it is a compact, convex subset of $\mathscr{M}_1(\Omega^\infty)$. So by Fact 4, K is the closed convex hull of its extreme points. What this literally means is that, for any $\mu \in K$, there is a sequence of measures $\rho_1, \rho_2, \dots \in \mathscr{M}_1([0,1])$, each of which is purely atomic with finite support, such that

$$\mu(\cdot) \;=\; \mathrm{w}-\lim_{n\to\infty} \int_0^1 \alpha_p^{\otimes\infty}(\cdot)\, d\rho_n(p)\,.$$

But this is not the type of representation one wants: one does not want to know merely that μ is a limit point of a sequence of mixtures of i.i.d. product measures; one wants to prove that μ is an i.i.d. product measure.

Fact 5. *The mapping from $\mathscr{M}_1([0,1])$ to $\mathscr{M}_1(\Omega^\infty)$, given by*

$$\rho \mapsto \int_0^1 \alpha_p^{\otimes\infty}(\cdot)\, d\rho(p)\,,$$

is a homeomorphism onto its image.

Before finishing the Hewitt and Savage proof of de Finetti's theorem, let us mention that there is an alternative, somewhat more constructive proof, which uses Doob's reversed martingale convergence theorem, and Kolmogorov's extension principle.

Let us state another fact, which we will need next lecture.

Fact 6. *For each $M, N \in \mathbb{Z}_{>0}$, with $M \leq N$, define $\phi_M^N : \Omega^N \to \Omega^M$ by $\phi_M^N(\sigma_1, \dots, \sigma_N) = (\sigma_1, \dots, \sigma_M)$. Suppose there is a sequence of measures $\mu_1, \mu_2, \dots$, such that $\mu_N \in \mathscr{M}_1(\Omega^N)$ for each $N \in \mathbb{Z}_{>0}$ which is consistent in the sense that*

$$\mu_N \circ (\phi_M^N)^{-1} \;=\; \mu_M$$

for each $M \leq N$. Then there is a unique measure $\mu \in \mathscr{M}_1(\Omega^\infty)$, such that $\mu_N = \mu \circ \phi_N^{-1}$ for each N.

Now for the proof of de Finetti's theorem. Suppose that $\rho_1, \rho_2, \dots$ is a sequence, as above. Then, since $\mathscr{M}_1([0,1])$ is compact, there is a convergent subsequence, which, without loss of generality, we may take to be the original sequence. By continuity,

$$\mu(\cdot) \;=\; \int_0^1 \alpha_p^{\otimes\infty}(\cdot)\, d\rho(p)\,,$$

where $\rho = \lim_{n\to\infty} \rho_n$. But by invertibility, there is no other choice for such a ρ than this one.

REFERENCES

[1] G. Choquet, *Lectures on analysis. Vol. I: Integration and topological vector spaces. Edited by J. Marsden, T. Lance and S. Gelbart*, W. A. Benjamin, Inc., New York-Amsterdam, 1969.

[2] N. Dunford and J. T. Schwartz, *Linear operators. Part I. General theory. With the assistance of William G. Bade and Robert G. Bartle. Reprint of the 1958 original*, Wiley Classics Library. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1988.

[3] R. M. Dudley, *Real analysis and probability. Revised reprint of the 1989 original*, Cambridge Studies in Advanced Mathematics, 74. Cambridge University Press, Cambridge, 2002.

[4] E. Hewitt and L. J. Savage. Symmetric measures on Cartesian products. *Trans. Amer. Math. Soc.* **80**, 470–501, 1955.

[5] J. L. Kelley. *General topology. Reprint of the 1955 edition [Van Nostrand, Toronto, Ont.]*, Graduate Texts in Mathematics, No. 27. Springer-Verlag, New York-Berlin, 1975.

[6] H. L. Royden, *Real analysis. Third edition*, Macmillan Publishing Company, New York, 1988.

[7] B. Simon, *The statistical mechanics of lattice gases. Vol. I*, Princeton Series in Physics. Princeton University Press, Princeton, NJ, 1993

MATHEMATICS DEPARTMENT, UNIVERSITY OF ROCHESTER, ROCHESTER, NY 14627
E-mail address: sstarr@math.rochester.edu