Measure concentration: Levy's Lemma

Lecture notes for Talk 6, on

Selected topics in Mathematical Physics: Quantum Information [3]

Manuel Gerken

December 3, 2013

Abstract

n-dimensional pure quantum states can be described as points on the surface of the 2n-dimensional unit sphere. In this talk we will observe some property of the sphere to gain knowledge of the concentration measure which will be interesting when discussing random pure quantum states. After a short introduction we will introduce concentration measure, discuss isoperimetric isotropy of the sphere and calculate the volume of the spherical cap. Afterwards we will be able to formulate Levy's Lemma, which gives an estimation of the probability distribution around the expectation value.

1 Levy's Lemma

Any pure quantum state $|\psi>$ has some complex coordinates $|\psi>=(z_1, z_2, ..., z_n)$ where $z_j \in \mathbb{C}$ for $j = 1, ..., n$. As we know the normalization of the vector gives $\sum_i |z_j|^2 = 1$. We can write $z_j = x_j + iy_j$ where $x_j, y_j \in \mathbb{R}$ so that $\sum_j^n x_j^2 + \sum_j^n y_j^2 = 1$. This shows, that we can describe every pure quantum-state $|\psi>$ as a vector on the 2n-dimensional unit sphere in $\mathbb{R}$.
As every point on the surface of a sphere $S^{(2n-1)}$ in $\mathbb{R}^{2n}$ describes a quantum states it can be interesting to examine some properties of the unit sphere.
We will soon see that some geometric properties of hyper spheres are very surprising and not trivial at all. For this let us consider a unit sphere S^{2n-1} in $\mathbb{R}^{2n}$. We will randomly choose a coordinate x_j and consider

$$S_\epsilon := \{x \in S^{(2n-1)} \mid d(x_j, 0) \leq \frac{\epsilon}{2}\}$$

where $d(x, y), \forall x, y \in S^{(2n-1)}$ is the angular norme on the sphere with $d(x, y) = \arccos\langle x, y\rangle$. The subset S_ϵ is an equator of width ϵ.
For an example considering the earth and choose the x_1 coordinate from south to north. The set S_ϵ will then be located around the equator of the earth. Choosing the x_2 coordinate from west to east, S_ϵ will be a strip including north and south pole.
With the normalized surface measure $\mu_{2n-1}(S^{(2n-1)}) = 1$ we can than show that

$$\mu_{2n-1}(S_\epsilon) \geq 1 - exp(-cn\epsilon^2)$$

where $c > 0$ is some constant. This means that almost all surface measure of the sphere is concentrated around the equator. This is not trivial at all, if you consider that the chosen x_i

coordinate is not special this has to be true for any other coordinate x_i' and every other equator on the sphere. If you now intersect two equator strips, their intersection will still have measure close to 1.
As we will see later this problem can be described more generally. We will be considering a Lipschitz continuous function $f : S^{2n-1} \to \mathbb{R}$ with the expectation value $\mathbb{E}_f$ with respect to the uniform measure on the sphere. We will show that randomly picked states are very close to $\mathbb{E}_f$ with a very high probability. This is formulated in Levy's Lemma:

Lemma 1 (Levy's Lemma)**:** *Let $f : S^{(2n-1)} \longrightarrow \mathbb{R}$ be Lipschitz-continuous with Lipschitz constant η, i.e.*

$$|f(x) - f(y)| \leq \eta \cdot \|x - y\|,$$

where $\|\cdot\| := \|\cdot\|_e$ denotes the Euclidean norm in the surrounding space $\mathbb{R}^{2n} \supset S^{(2n-1)}$. Drawing a point $x \in S^{(2n-1)}$ at random with respect to the uniform measure on the sphere yields

$$Prob\{|f(x) - \mathbb{E}_f| \geq \epsilon\} \leq 2exp(-\frac{n\epsilon^2}{9\pi^3\eta^2})$$

for all $\epsilon \leq 0$.

We will derive and formulate another version of Levy's Lemma in the end of this talk. To proof it we will take a detour discussing concentration measure and the Lemma of Isoperimetric inequality. A physical interpretation of the result will be given in the last section of this talk.

2 Concentration Measure

As the name says, concentration measure is a way of measuring concentration on a metric measure space. In the introduction we already claimed that most of the surface measure is concentrated in S_ϵ. To make this claim comparable we need to observe the concentration measure of this set. In the following part we will have a triplet (X, d, μ) where X is a separable complete metric measure space,
$d : X \times X \longrightarrow \mathbb{R}$ is a function that satisfies the metric axioms, [i.e. $d(x, y) \geq 0$, $d(x, y) = d(y, x)$ and $d(x, y) \leq d(x, z) + d(z, y)$, $\forall x, y, z \in X$] and μ is a finite Borel measure. We define the median M_f

Definition 1 (Median)**:** *Let X be a metric measure space, and $f : X \longrightarrow \mathbb{R}$ a continuous function. A median M_f is defined by*

$$\mu\{x \in X \mid f(x) \leq M_f\} = \frac{1}{2} \tag{1}$$

The median is a very intuitive value. Other than the expectation value, the median can be seen on first sight in most cases. For example let A be a set with $A := \{2, 5, 7, 43, 78, 94, 154, 256, 472, 715\}$, μ the relative counting measure and $f(x) = x$ with $x \in A$ a Lipschitz continuous function in A, then the median is $M(A) = 78$ whereat the expectation value has to be calculated. The median M_f is not necessarily unique, even though in most cases it is. It is Important to note that $\mathbb{E}f \neq M_f$.
Now we can introduce the *concentration function* α_X over a metric measure space X with $\mu(X) = 1$. $\alpha_X(S)$ captures the concentration properties of the set S. If it is small, we have strong concentration of measure.

Definition 2 (Concentration function)**:** *Let X be a metric measure space. For every $\epsilon > 0$, define the concentration function*

$$\alpha_X(\epsilon) := \sup\{\ \mu(X \backslash N_\epsilon(S)) \mid S \text{ is measurable and } \mu(S) = \frac{1}{2}\}$$

where $N_\epsilon(S)$ is the ϵ-neighbourhood of S:

$$N_\epsilon(S) = \{\ x \in X \mid \exists s \in S\ :\ d(s,x) < \epsilon\}$$

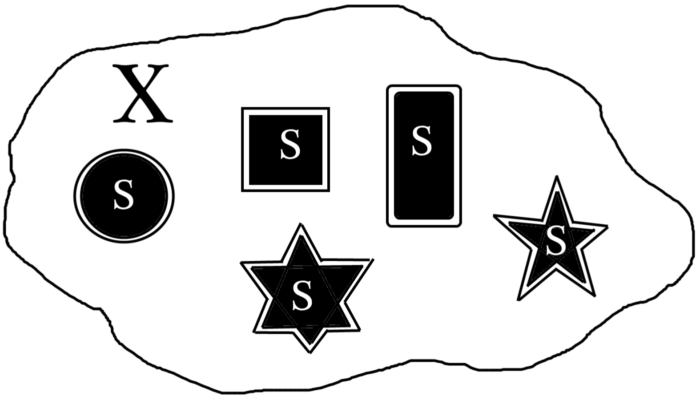

Figure 1: A metric measure space X, possible subsets S with $\mu(S) = const.$ declared as the black areas and $N_\epsilon(S)$ as the black areas including the white border with constant ϵ. For the different subsets S the ϵ-neighbourhood is not equal. The circle has the smallest ϵ-neighbourhood, thus $\alpha_X(\epsilon) = \mu(X \backslash N_\epsilon(S_{circle}))$. (note that $\mu(S) \neq \frac{1}{2}$ for the example)

In Fig.1 is an example to better understand concentration measure. On the metric measure space X we can observe different subsets S with $\mu(S) = \frac{1}{2}$. Depending on the form of the subset (in the example we have different forms) the ϵ-neighbourhood has a different measure, $\mu(N_\epsilon(S_{circle})) \neq \mu(N_\epsilon(S_{square})) \neq \mu(N_\epsilon(S_{star})) \neq ...$ $\alpha_\epsilon(X)$ is then the white area measure of X minus the measure of the smallest ϵ-neighbourhood.

$$\alpha_X(\epsilon) = \mu(X) - \inf\{\mu(N_\epsilon(S))\}$$

where $\mu(X) = 1$. If $\alpha_X(\epsilon)$ is small, the space has a strong concentration of measure.
To fully understand this concentration function and to visualize it we will need another notation of $\alpha_X(\epsilon)$.

Lemma 2: *Let X be a metric measure space and $f : X \longrightarrow \mathbb{R}$ a Lipschitz-continuous function with constant 1, then*

$$\mu\{x \in X \mid f(x) \geq M_f + \epsilon\} \leq \alpha_x(\epsilon) \tag{2}$$

Proof. Let $S := \{x | f(x) \leq M_f\}$ then we know from the median Definition.1, that $\mu(S) = \frac{1}{2}$ and thus $\mu(X \backslash N_\epsilon(S)) \leq \alpha_X(\epsilon)$. If we consider a subset $A \subseteq X$ with $f(a) \geq M_f + \epsilon$, $\forall a \in A$, then we know because f is Lipschitz continuous function with constant 1, that $a \notin N_\epsilon(S), \forall a \in A$.
Now we can conclude that$\{a \in X | f(a) \geq M_f + \epsilon\} \subseteq X \backslash N_\epsilon(S)$.
Thus we can say: $\mu\{x \in X | f(x) \geq M_f + \epsilon\} \leq \mu(X \backslash N_\epsilon(S)) \leq \alpha_X(\epsilon)$. □

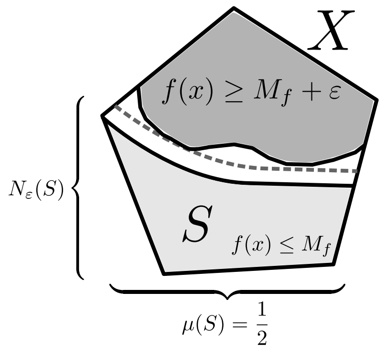

Figure 2: Metric measure space X with subsets S and $N_X(\epsilon)$

One could even show equality for the above Lemma, but we do not need this for our Lecture. Fig.2 shows the properties of $\alpha_X(\epsilon)$ as explained in the Lemma.2.
Consider any function f which is Lipschitz continuous with constant (less than or equal) to 1. Any subset S of points $x \in X$ with the property $f(x) \leq M_f$ has by definition of the median the measure $\mu(S) = \frac{1}{2}$. The whole area under the dashed line implying the grey S area indicates the $\epsilon - neighborhood$ $N_\epsilon(S)$. Because of the Lipschitz continuity, all points $x \in N_\epsilon(S)$ satisfy $f(x) < M_f + \epsilon$. So the darker grey area with the set of points $x \in X$ that satisfy $f(x) \geq M_f + \epsilon$ must be a subset of the complement $X \backslash N_\epsilon(S)$.
For a Lipschitz continuous function f with Lipschitz constant η we just have to change the ϵ to $\epsilon \rightarrow \epsilon\eta = \epsilon'$, because of the Lipschitz continuity $|f(x) - f(y)| \leq \eta \|x - y\| \leq \epsilon\eta$.
The relative measure μ has all the properties of a Probability measure, thus we can do the first step towards Levy's Lemma by saying, that measure concentration means that the dark shaded region should be as small as possible. Since it is upper-bounded by $\mu(X \backslash N_\epsilon(S))$, which is in turn upper-bounded by $\alpha_X(\epsilon)$, (the best possible upper-bound) we can say that:

$$Prob\{f(x) \geq M_f + \epsilon'\} \leq \alpha_X(\frac{\epsilon'}{\eta}) \tag{3}$$

3 Isoperimetric inequality for the sphere

Since in our case we are interested in a particular metric measure space we will now be more specific about what we are looking at.
The metric measure space will now be $X = S^{(2n-1)}$, the sphere of dimension (2n-1). The dimension has been discussed in the introduction of Levy's Lemma. We are looking at a $\mathbb{C}^n$ dimensional pure quantum states, being represented on the surface of a sphere in $\mathbb{R}^{2n}$. The metric will now be $d(x, y) = \arccos\langle x, y\rangle \; where \; x, y \in S^{(2n-1)}$, which is the angle between x and y and the measure will be the uniform geometric measure.
The next step will now be to compute $\alpha_X(\epsilon)$ for our particular problem. For this we will use the Lemma of the Isoperimetric inequality for the sphere. We will need the definition of a *spherical cap* $B(a, r)$ which is the set of points $x \in X$ with $d(a, x) \leq r$ as it can be seen in Fig.3.

Lemma 3 (Isoperimetric inequality for the sphere)**:** *Let $A \subseteq S^{(2n-1)}$ be a closed subset of the sphere, and let $B = B(a,r) \subset S^{(2n-1)}$ be a spherical cap around any point $a \in S^{(2n-1)}$, where the radius r is chosen such that $\mu(B) = \mu(A)$. Then*

$$\mu(N_\epsilon(A)) \geq \mu(N_\epsilon(B)) \tag{4}$$

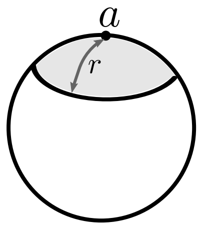

Figure 3: The spherical cap $B(a,r)$ on $S^{(2n-1)} \subset \mathbb{R}^{2n}$solves the isoperimetric problem for the sphere

This will not be proven. As it might be intuitive on a three dimensional sphere it is not trivial in more dimensions. The proof is very complicated and can be found in many different books. The isoperimetric inequality of the sphere says, that the spherical cap is the best possible form on a sphere with respect to the border/area proportion. If we remember the proposed forms in Fig.1 we can now say what $\alpha_X(\epsilon)$ is.

4 Computing $\alpha_X(\epsilon)$

This Lemma of the Isoperimetric inequality shows us, that the measure on the unit sphere with the smallest border or ϵ-expansion, is the cap $B(a,r)$ as shown in Fig.3. To calculate $\alpha_{S^{(2n-1)}}(\epsilon)$ we now have to calculate $B(a,r)$.

$$\alpha_{S^{(2n-1)}}(\epsilon) = 1 - \inf\{\mu(N_\epsilon(S))\} = 1 - \mu(N_\epsilon(B(a, \frac{\pi}{2}))) = 1 - \mu(B(a, \frac{\pi}{2} + \epsilon)) \tag{5}$$

We will observe the normalized volume element on the unit sphere $S^{(2n-1)}$[1].

For the normalized measure μ of the sphere we know that $A(r) := \mu(B(a,r))$ is the surface of dimension $\dim(\mu(A)) = 2n-1$. We will substitute $2n-1 = m$ for simplification. For the surface we have

$$A(\phi) = s_m^{-1} \int_0^\phi \sin^{(m-1)}(\phi')d\phi' \tag{6}$$

where $s_m = \int_0^\pi \sin^{m-1}(\phi')d\phi'$. To calculate $\alpha_{S^m}(\epsilon)$ we know $\alpha_{S^m}(\epsilon) = 1 - A(\frac{\pi}{2} + \epsilon)$ out of equation.(5) with:

$$\begin{aligned} 1 - s_m^{-1} \int_0^{\frac{\pi}{2}+\epsilon} \sin^{(m-1)}(\phi)d\phi &= s_m^{-1} \int_{\frac{\pi}{2}+\epsilon}^{\pi} \sin^{(m-1)}(\phi)d\phi \\ &= s_m^{-1} \int_\epsilon^{\frac{\pi}{2}} \cos^{(m-1)}(\phi)d\phi \end{aligned}$$

Using the elementary inequality $\cos(u) \leq \exp(-\frac{u^2}{2})$ for $0 \leq u \leq \frac{\pi}{2}$ we will replace the cos. By substituting $\phi = \frac{\tau}{\sqrt{m-1}}$ we get: $d\phi = \frac{d\tau}{\sqrt{m-1}} \Rightarrow d\tau = d\tau\sqrt{m-1}$ and for the constraints

$\frac{\pi}{2} \to \frac{\pi}{2}\sqrt{m-1}$, and $\epsilon \to \epsilon\sqrt{m-1}$:

$$\begin{aligned}
s_m^{-1}\int_{\epsilon}^{\frac{\pi}{2}} \cos^{(m-1)}(\phi)d\phi &= \frac{s_m^{-1}}{\sqrt{m-1}}\int_{\epsilon\sqrt{m-1}}^{\frac{\pi}{2}\sqrt{m-1}} \cos^{(m-1)}(\frac{\tau}{\sqrt{m-1}})d\tau \\
&\leq \frac{s_m^{-1}}{\sqrt{m-1}}\int_{\epsilon\sqrt{m-1}}^{\infty} (\exp(-\frac{1}{2}(\frac{\tau}{\sqrt{m-1}})^2)^{(m-1)}d\tau \\
&\leq \frac{s_m^{-1}}{\sqrt{m-1}}\int_{\epsilon\sqrt{m-1}}^{\infty} \exp(-\frac{\tau^2}{2})d\tau \\
&\leq s_m^{-1}\frac{\sqrt{\pi}}{\sqrt{2(m-1)}}\exp(-(m-1)\frac{\epsilon^2}{2})
\end{aligned}$$

We can replace the upper constraint $\frac{\pi}{2}\sqrt{m-1}$ by infinity when we replace the equality by a less-equal, since $exp(-x^2) \geq 0,\ \forall x \in \mathbb{R}$. The integral has been solved with the solution of the Error-function $\Phi(\epsilon)$. We know that $1 - \Phi(\epsilon) \leq \frac{1}{2}\exp(-\frac{\epsilon^2}{2})\ for\ \epsilon \geq 0$.
Left is the value of s_m^{-1}. We know from partial integration that: $s_m = [\frac{m-2}{m-1}]s_{m-2}$ for $m \geq 2$. Out of $s_m\sqrt{m-1} = \sqrt{m-1}s_{m-2}$ follows $\sqrt{m-1}s_m \geq 2 \Rightarrow s_m^{-1} \leq \frac{\sqrt{m-1}}{2}$ and since $\sqrt{\frac{\pi}{8}} \leq 1$ we have:

$$\alpha_X(\epsilon) \leq \exp(-(2n-2)\frac{\epsilon^2}{2}) \tag{7}$$

5 Final version of Levy's Lemma

By now we know that $\mu\{f(x) \geq M_f + \epsilon\} \leq \alpha_X(\epsilon)$ for $f(x)$ a Lipschitz continuous function with constant less or equal to 1.
What we actually want is the probability i.e. the relative measure

$$\mu\{|f(x) - M_f| \geq \epsilon\} \leq \alpha'_X(\epsilon) \tag{8}$$

i.e.

$$\mu\{f(x) \geq (M_f + \epsilon)\ or\ f(x) \leq (M_F - \epsilon)\} \leq \alpha'_X(\epsilon) \tag{9}$$

For this we need to evaluate $\mu\{f'(x) \leq M_f - \epsilon\}$.
Since our function $f(x)$ is Lipschitz continuous we can know $g(x) = -f(x)$. The rest of the calculation will be analogue to the proof of Lemma 2 $\Rightarrow \mu\{g(x) \leq M_f - \epsilon\} \leq \alpha_X(\epsilon)$. $\Rightarrow$

$$\mu\{|f(x) - M_f| \geq \epsilon\} \leq 2\alpha_X(\epsilon) \tag{10}$$

The next thing to check will be the norm. As we used the angular distance $d_a(x,y) = \arccos\langle x,y\rangle$ we will have to change it to the Euclidean norm $d_e(x,y) = |x-y|$ and thus estimate a euclidean $\alpha_X^e(\epsilon)$. This appears to be quite easy since $d_e(x,y) \leq d_a(x,y)$ as we can see in Fig.4. Thus we also know that the $\epsilon - neighborhood$ of the Euclidean norm is bigger then the one of the angular distance norm. That is why $N_\epsilon^a(S) \subseteq N_\epsilon^e(S) \Rightarrow \mu\{N_\epsilon^a(S)\} \leq \mu\{N_\epsilon^e(S)\}$.
For a Lipschitz continuous function $f(x)$ for the Euclidean distance norm we have $|f(x)-f(y)| \leq \eta\|x-y\|_e \leq \eta\|x-y\|_a$. This way we can choose a function that is Lipschitz continuous with respect to the Euclidean norm, but still use the calculated values of the angular distance norm. If we observe the upper bound calculated in section 4 we can now say:

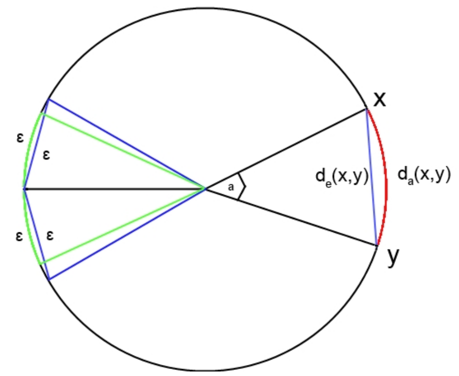

Figure 4: The angular distance $d_a(x, y)$ compared to the Euclidean distance $d_e(x, y)$. Since $d_e(x, y) \leq d_a(x, y)$ the $\epsilon - neighborhood$ of the Euclidean norm is bigger.

$$\begin{aligned}
\alpha_X^e(\epsilon) &= 1 - \inf\{\mu(\underbrace{N_\epsilon^e(S)}_{\geq N_\epsilon^a(S)}) \mid S \ measurable, \ \mu(S) = \frac{1}{2}\} \\
&\leq 1 - \inf\{\mu(N_\epsilon^a(S)) \mid S \ measurable, \ \mu(S) = \frac{1}{2}\} \\
&\leq \alpha_X^a(\epsilon)
\end{aligned}$$

This way we can use $\alpha_X^a(\epsilon)$ as an upper bound for the probability measure in the Lemma.
Another important point is the ϵ as we had some modifications in section 2 where we said $\epsilon \to \frac{\epsilon}{\eta}$ where η was the Lipschitz constant.
In our first definition of Levy's Lemma we had the expectation value $\mathbb{E}f$ instead of the median M_f. These two values are not the same. An inequality can be shown which brings the missing factors to the exponential function when changing median and expectation value.
Now we are ready to state the final version of Levy's Lemma:

Lemma 4 (Levy's Lemma)**:** *Let $f : S^{(2n-1)} \longrightarrow \mathbb{R}$ be Lipschitz-continuous with Lipschitz constant η, i.e.*

$$|f(x) - f(y)| \leq \eta \cdot \|x - y\|_e, \tag{11}$$

where $\|\cdot\|_e := \|\cdot\|_2$ denotes the Euclidean norm in the surrounding space $\mathbb{R}^{2n} \supset S^{(2n-1)}$. Drawing a point $x \in S^{(2n-1)}$ at random with respect to the uniform measure on the sphere yields

$$Prob\{|f(x) - M_f| \geq \epsilon\} \leq 2exp(-n\frac{\epsilon^2}{\eta^2}) \tag{12}$$

for all $\epsilon \leq 0$.

6 Physical Interpretation

The conclusion of the last talk about 'Averages over the unitary group' was the expectation value of random pure quantum states in two coupled systems A and B with $|A| \ll |B|$. We derived that $\mathbb{E}_\psi Tr(\rho_A^2) \approx \frac{1}{|A|}$, i.e. the expectation value of a randomly picked pure quantum state is the most entangled state[2].
Now we showed that almost all quantum states are almost maximally entangled.

This is analogue to the partition function in statistical physics. In the following talks we will hear more about physical interpretation of these results and the direct connection to statistical mechanics.

References

[1] M. Ledoux. *The concentration of measure phenomenon, Mathematical Suveys and Monographys 89, AMS, USA*.

[2] Isaac L. Chung Michael A. Nielsen. *Quaantum Computation and Quantum Information.* Elsevier (2010).

[3] Markus Müller. Random quantum states, measure concentration, and the additivity conjecture.