These notes are based on the work of Jean Bourgain titled *Periodic nonlinear Schrödinger equation and invariant measures*. In particular of great help was the exposition by Oh, Sosoe, and Tolomeo in their work titled *Optimal integrability threshold for Gibbs measures associated with focusing NLS on the torus*.

1 Introduction

We will attempt to answer the following question

Question 1. *Define the energy function*

$$H(u) = \frac{1}{2}\int |\nabla u|^2\,dx - \frac{1}{p}\int |u|^p\,dx. \tag{1}$$

Is it possible to sample a random function $u : \mathbb{S}^1 \to \mathbb{R}$ such that $\mathbb{P}[u] \propto e^{-\beta H(u)}$?

Note that we use ∇ to denote the derivative ∂_x with respect to the (single) space variable, simply to be suggestive about how things may be extended to higher dimensions. Similarly, we will use Δ to denote ∂_x^2.

1.1 Motivation

The reason for asking Question 1 comes from the consideration of the *nonlinear Schrödinger equation*

$$i\partial_t u + \Delta u + |u|^{p-2}u = 0.$$

Note that the flow of this equation preserves the energy (1) as can be seen by the following short calculation. First note that $\partial_t|a|^2 = 2\,\mathrm{Re}(a\partial_t\bar{a})$, so we have

$$\begin{aligned}
\partial_t H(u) &= \int \frac{1}{2}\partial_t|\nabla u|^2 - \frac{1}{p}\partial_t|u|^p \;\; dx \\
&= \int \mathrm{Re}\big(\nabla u\nabla\partial_t\bar{u} - |u|^{p-2}u\partial_t\bar{u}\big)\,dx \\
&= -\,\mathrm{Re}\int (\Delta u + |u|^{p-2}u)\partial_t\bar{u}\,dx \\
&= \mathrm{Re}\left(i\int |\partial_t u|^2\,dx\right) = 0.
\end{aligned}$$

So in principle, the Gibbs measure with weight proportional to $e^{-\beta H(u)}$ should be an invariant measure for the flow. If this is a *probability measure*, then we can use some ergodic theory considerations to derive useful results about the PDE. For instance, such a measure would give us a class of functions for which we have *global* well-posedness, since by sampling from the Gibbs measure we will, with probability 1, get a function for which the flow is well-posed by e.g. the Poincaré recurrence theorem.

One might ask about the particular form of the Gibbs measure $e^{-\beta H(u)}$. After all, wouldn't any function of the energy give a measure with the same property? The answer to that is yes in principle, but for reasons we'll see later, it is often easier to make sense of things when the energy is in the exponent, especially if the energy has a term of the form $\|\nabla u\|_2^2$, as this will lead us to an analysis in terms of a *Gaussian measure*. In addition, the exponential form shows up naturally in statistical mechanics, where β plays the role of an inverse temperature, and we imagine our system in contact with an environment at that temperature which allows for energy exchange. However, since we are considering deterministic dynamics with no energy transfer (as energy is conserved), this doesn't matter so much, and for our purposes we can just ignore β, so we'll set it to 1.

Now, as stated, Question 1 is false as the energy decreases as u gets larger, so simply shifting it towards $\pm\infty$ will show that there cannot be such a measure as it would have to have infinite mass at least. Even if we restrict u to have total integral 0, say, the energy can be exchanged between the two summands and thus be preserved while both summands go off to infinity; this forms a *soliton*.

Indeed, in simulations you can observe soliton-like spikes forming, moving and interacting. However, since the total mass (squared L^2 norm) of u is also preserved under the flow, the spikes will not grow to infinity in many instances. For this reason, the Gibbs measure formula $e^{-H(u)}$ must account for functions of any mass, and thus will not likely be a probability measure. Here is a quick proof that the mass is preserved under the flow:

$$\begin{aligned}\partial_t \int |u|^2\,dx &= \int 2\operatorname{Re}(u\partial_t\bar{u})\,dx \\ &= 2\operatorname{Re}\left(-i\int\ u\Delta\bar{u} + u|u|^{p-2}\bar{u}\ \ dx\right) \\ &= 2\operatorname{Re}\left(-i\int\ -|\nabla u|^2 + |u|^p\ \ dx\right) = 0.\end{aligned}$$

So one way to try and salvage Question 1 is to add a mass cutoff, and we arrive at

Question 2. *Does the following Gibbs (probability) measure exist?*

$$\mathbb{P}[u] \propto \exp\left(-\frac{1}{2}\|\nabla u\|_2^2 + \frac{1}{p}\|u\|_p^p\right)\cdot \mathbf{1}_{\{\|u\|_2\le M\}}$$

1.2 Relevance of the GNS inequality

Recall the following version of the Gagliardo-Nirenberg-Sobolev inequality (for $d=1$):

$$\|u\|_p^p \le C\|\nabla u\|_2^{\frac{p-2}{2}}\|u\|_2^{\frac{p+2}{2}}.$$

In fact, we will only need a slightly weaker version which I guess is sometimes called Bernstein's inequality:

$$\begin{aligned}\|u\|_p &\le C\|\nabla u\|_2^{\frac{1}{2}-\frac{1}{p}}\|u\|_2^{\frac{1}{2}+\frac{1}{p}} \\ &\le C\|\nabla u\|_2^{\frac{1}{2}-\frac{1}{p}}\|u\|_2.\end{aligned}$$

For some intuition, if $\|u\|_p$ is large then it might be the case that $\|u\|_2$ is also large; this would happen if the whole function was large everywhere. But if $\|u\|_p$ is large and $\|u\|_2$ is small, then u must be very spiky since the contribution should be coming from some peaks which are much more relevant in the calculation of the p-norm since $p>2$. In this case, the $\|\nabla u\|_2$ term takes over. The precise powers can be derived via scaling.

We will see how to use this inequality rigorously later. For now, let us see how this can be used to understand Question 2. For the probability measure there to make sense, it should be normalizable, i.e. we should have

$$\int \exp\left(-\frac{1}{2}\|\nabla u\|_2^2 + \frac{1}{p}\|u\|_p^p\right)\cdot \mathbf{1}_{\{\|u\|_2\le M\}}\,du < \infty.$$

Note that I'm not saying what the reference measure on function is just yet. In any case, heuristically, by the GNS inequality we see that

$$\int_{\|u\|_2\le M} \exp\left(-\frac{1}{2}\|\nabla u\|_2^2 + \frac{1}{p}\|u\|_p^p\right)du \le \int_{\|u\|_2\le M}\exp\left(-\frac{1}{2}\|\nabla u\|_2^2 + \frac{C}{p}M^{\frac{p+2}{2}}\|\nabla u\|^{\frac{p-2}{2}}\right)du.$$

Now, if $\frac{p-2}{2}<2$, i.e. $p<6$, then the expression in the exponent goes to $-\infty$ as $\|\nabla u\|_2\to\infty$, which hints that the measure should maybe be normalizable. Additionally, if $p=6$ and M is small enough we should get the same result. However, if $p>6$ then the measure should not be normalizable, and thus the probability measure of Question 2 should not exist.

In these notes, we will prove the first statement, that the measure is normalizable for $p<6$. First, however, we'll have to make sense of the question, which is not quite well-formed yet. This is because there is no analog of Lebesgue measure on all functions, so thinking of a probability density on all functions just doesn't make sense. But we will make sense of this using a canonical measure on the space of functions, called the Gaussian free field.

2 The Gaussian free field

Before attempting to answer Question 2, we should make sense of a simpler question: what if we just ignore the problematic $\|u\|_p^p$ term in the exponent?

Question 3. *What do we mean by the following probability measure?*

$$\mathbb{P}[u] \propto \exp\left(-\frac{1}{2}\|\nabla u\|_2^2\right)$$

We can actually answer this question in a somewhat general way, i.e. in a wide range of domains and in higher dimensions (although the rest of the talk outside of this section is restricted to $d = 1$); the resulting object is called the *Gaussian free field.*

2.1 A primer on multivariate (complex) Gaussians

Recall that a $\mathbb{R}$-valued random variable X is Gaussian with variance σ^2 if

$$\mathbb{P}[X = x] \propto \exp\left(-\frac{x^2}{2\sigma^2}\right).$$

Note that this has the same distribution as σg, where g is a standard Gaussian (variance 1).

More generally, for any positive semidefinite $d \times d$ matrix Σ, the $\mathbb{R}^d$-valued random vector X is Gaussian with *covariance matrix* Σ if

$$\mathbb{P}[X = x] \propto \exp\left(-\frac{1}{2}\langle x, \Sigma^{-1}x\rangle\right).$$

This means that $\Sigma_{i,j} = \mathrm{Cov}[X_i, X_j]$. In particular, if Σ is not diagonal, the coordinates of the Gaussian vector will not be independent. So let's diagonalize Σ, writing ϕ_j for the orthonormal basis of eigenvectors and σ_j^2 for the corresponding eigenvalues. Let's assume that all of these are strictly positive for now. Then letting $x = \sum_{j=1}^d \alpha_j \phi_j$, we find that

$$\mathbb{P}[X = x] \propto \exp\left(-\frac{1}{2}\sum_{j=1}^d \frac{1}{\sigma_j^2}\alpha_j^2\right) = \prod_{j=1}^d \exp\left(-\frac{\alpha_j^2}{2\sigma_j^2}\right).$$

In other words, the coefficients in the basis of eigenvectors for Σ are independent Gaussians with variance σ_j^2. This means that we may write down a vector with the same distribution by setting

$$X = \sum_{j=1}^d \sigma_j g_j \phi_j,$$

where g_j are independent standard normal random variables.

All of this generalizes to complex-valued Gaussians easily; if we have a complex number Z with

$$\mathbb{P}[Z = z] = \exp\left(-\frac{|z|^2}{2\sigma^2}\right),$$

then the real and imaginary parts of Z are independent Gaussains, each with variance σ^2.

2.2 Diagonalizing the Laplacian

Let's write $\|\nabla u\|_2^2 = \langle u, -\Delta u\rangle$; in some sense, the measure we're after in Question 3 is the Gaussian with covariance matrix $(-\Delta)^{-1}$. As we saw in the previous section, it's better to try and diagonalize this operator, which is almost equivalent to diagonalizing $-\Delta$. Let's write $\{\phi_n\}$ for the eigenfunctions of $-\Delta$, with eigenvalues $\{\lambda_n\}$. Note that (assuming our domain is connected), one eigenvalue (let's say λ_0) is 0 and the rest are positive. Then, if we write $u = \sum_n \alpha_n \phi_n$, we have

$$\exp\left(-\frac{1}{2}\langle u, -\Delta u\rangle\right) = \exp\left(-\frac{1}{2}\sum_n \lambda_n |\alpha_n|^2\right).$$

In other words, under the measure in Question 3, the coefficients α_n behave like independent complex Gaussians with variance $\frac{1}{\lambda_n}$.

Note that we need to be a bit careful about the term corresponding to the zero eigenvalue; that coefficient should have *infinite* if we don't make any other assumptions. This is consistent with the fact that $\|\nabla u\|_2^2$ is unchanged when we shift u by a constant, so without any other assumptions, the GFF should be invariant under such shifts, but then it clearly can't be a probability measure. For this reason, the GFF must be centered in some way; for the current presentation, we center by assuming that $\int u\,dx = 0$, which means that the coefficient for the zero eigenvector (which is constant) is 0.

In other words, the following random variable should have the distribution of the GFF (recall $\lambda_0 = 0$):

$$u = \sum_{n \neq 0} \frac{g_n}{\sqrt{\lambda_n}} \phi_n,$$

where g_n are i.i.d. standard complex Gaussians. Note that in many situations, this is not differentiable, or maybe not even a function. So it doesn't quite make sense to calculate $\|\nabla u\|_2^2$ directly, but this can be made sense of and indeed this is the correct object.

2.3 The case of one dimension

Let's suppose that $\mathbb{S}^1$ has length 2π. This means that the eigenfunctions of $-\Delta$ are e^{inx} for $n \in \mathbb{Z}$, with eigenvalue n^2. So our GFF is the random function

$$u = \sum_{n \neq 0} \frac{g_n}{n} e^{inx}, \tag{2}$$

where g_n are i.i.d. standard complex Gaussians. This is also called a *Brownian bridge*, and it is the same thing as a Brownian motion conditioned to end where it started and conditioned to have mean zero. It may also be thought of as a random version of the Weierstrass function, which is the classic example of a function which is continuous but not differentiable. In fact, the Hölder regularity of u is $\frac{1}{2}-$, and in higher dimensions it is worse: $\frac{2-d}{2}-$, meaning that it is not even a function when $d \geq 2$.

2.4 Tilting away from the GFF

With this explicit representation for the GFF, we may return to Question 2. The measure written there should be a *tilt* away from the GFF, in the sense that the *Radon-Nikodym derivative* of that measure with respect to the GFF is

$$\exp\left(\frac{1}{p}\|u\|_p^p\right) \cdot \mathbf{1}_{\{\|u\|_2 \leq M\}}.$$

However, we need to check that the above function is integrable with respect to the GFF, so that the measure with that Radon-Nikodym derivative is normalizable to a probability measure. In other words, we have rephrased our problem as follows.

Question 4. *For which values of p and M is it true that*

$$Z_{p,M} = \mathbb{E}\left[\exp\left(\frac{1}{p}\|u\|_p^p\right) \cdot \mathbf{1}_{\{\|u\|_2 \leq M\}}\right] < \infty, \tag{3}$$

where u is the GFF defined in (2)?

Note that $Z_{p,M}$ is called the *partition function* for the Gibbs measure, if it is finite.

3 Bourgain's argument

In this section we will prove that if $2 < p < 6$, then (3) holds for all M, and if $p = 6$ then it holds for all small enough M. The proof will go by decomposing u across scales of frequencies, and bounding the p-norm of each scale separately. We'll use the GNS/Bernstein inequality to reduce the p-norm to the 2-norm of the function and its derivative ∇u. On each scale, we can bound $\|\nabla u\|_2$ appropriately by the frequency, and $\|u\|_2$ is easily accessible by the sum representation (2).

3.1 Rewriting in terms of a tail bound

First, by rewriting the expectation as an integral over t and changing variables $t = e^{\frac{1}{p}\lambda^p}$, we find that

$$\begin{aligned} Z_{p,M} &= \int_0^\infty \mathbb{P}\left[e^{\frac{1}{p}\|u\|_p^p} > t, \|u\|_2 \le M\right] dt \\ &= \int_0^\infty \mathbb{P}[\|u\|_p > \lambda, \|u\|_2 \le M] \cdot e^{\frac{1}{p}\lambda^p} \lambda^{p-1}\, d\lambda. \end{aligned}$$

Thus to show that $Z_{p,M} < \infty$ it suffices to show that there exists $C > 0$ and $c > \frac{1}{p}$ such that

$$\mathbb{P}[\|u\|_p > \lambda, \|u\|_2 \le M] \le Ce^{-c\lambda^p}$$

for all large enough λ.

3.2 Dividing into frequency scales

For $j \ge 0$, let us set

$$u_j = \sum_{2^j \le |n| < 2^{j+1}} \frac{g_n}{n} e^{inx}, \qquad u_{<k} = \sum_{j<k} u_j, \quad \text{and} \quad u_{\ge k} = \sum_{j \ge k} u_j.$$

Note that $\|u\|_p \le \|u_{<k}\|_p + \|u_{\ge k}\|_p$, so it suffices to get appropriate bounds on

$$\mathbb{P}\left[\|u_{<k}\|_p > \frac{\lambda}{2}, \|u\|_2 \le M\right] \qquad \text{and} \qquad \mathbb{P}\left[\|u_{\ge k}\|_p > \frac{\lambda}{2}, \|u\|_2 \le M\right]$$

We will ignore the $\|u\|_2 \le M$ restriction for the high-frequency part of u, i.e. $u_{\ge k}$, but we will keep it in for the low-frequency part, i.e. $u_{<k}$. The critical value of k where we need to split will become apparent shortly.

3.3 The low-frequency part

Note first that anything with only frequency components with frequency $\le n$ has derivative at most $O(n)$ as well. So let us apply GNS/Bernstein to $u_{<k}$ as follows, using the bound $\|u_{<k}\| \le \|u\|_2 \le M$:

$$\|u_{<k}\|_p \le C 2^{k(\frac{1}{2} - \frac{1}{p})} \|u_{<k}\|_2 \le C 2^{k(\frac{1}{2} - \frac{1}{p})} M.$$

Thus, if we set

$$k = \log_2\left(\left(\frac{\lambda}{2CM}\right)^{\frac{2p}{p-2}}\right), \tag{4}$$

then we find that $\|u_{<k}\|_p \le \frac{\lambda}{2}$ deterministically. Keep in mind that this means that k goes to ∞ as λ goes to ∞. Note also that the critical *frequencies* (not scales) are polynomial in terms of λ.

3.4 The high-frequency part

By the above work, it suffices to show that there exist $C > 0$ and $c > \frac{1}{p}$ such that for all λ large enough we have

$$\mathbb{P}\left[\|u_{\geq k}\|_p > \frac{\lambda}{2}\right] \leq Ce^{-c\lambda^p}$$

where $k = k(\lambda)$ is as in (4). Let us choose later a sequence $\{\lambda_j\}_{j=k}^{\infty}$ such that $\sum_{j=k}^{\infty} \lambda_j = \frac{\lambda}{2}$. Then we have

$$\mathbb{P}\left[\|u_{\geq k} > \frac{\lambda}{2}\right] \leq \sum_{j=k}^{\infty} \mathbb{P}[\|u_j\|_p > \lambda_j].$$

Now, using GNS/Bernstein again, we have

$$\|u_j\|_p \leq C2^{j(\frac{1}{2}-\frac{1}{p})}\|u_j\|_2,$$

which means that

$$\begin{aligned}\mathbb{P}[\|u_j\|_p > \lambda_j] &\leq \mathbb{P}\left[\|u_j\|_2 > \frac{1}{C}\lambda_j 2^{j(-\frac{1}{2}+\frac{1}{p})}\right] \\ &= \mathbb{P}\left[\sum_{2^j \leq |n| \leq 2^{j+1}} \frac{|g_n|^2}{n^2} > \frac{1}{C^2}\lambda_j^2 2^{j(-1+\frac{2}{p})}\right] \\ &\leq \mathbb{P}\left[\sum_{2^j \leq |n| \leq 2^{j+1}} |g_n|^2 > \frac{1}{C^2}\lambda_j^2 2^{j(1+\frac{2}{p})}\right]. \end{aligned} \tag{5}$$

Now inside we have a sum over many independent random variables, meaning that the central limit theorem (or more specifically a Gaussian tail bound) can be applied. The mean of this sum is 2^j and the standard deviation is of order $2^{j/2}$, so a standard large-deviations type result is that we get an exponentially small tail probability for the sum to be larger than some constant multiple of 2^j.

In other words, in order to get a small bound it would suffice to have

$$\lambda_j^2 2^{j(1+\frac{2}{p})} \gg 2^j \qquad \text{i.e.} \qquad \lambda_j \gg 2^{-j/p}, \tag{6}$$

In order to have (6) and still have the sum be $\frac{\lambda}{2}$, we can choose λ_j decaying exponentially with rate $< 2^{-\frac{1}{p}}$ with appropriate constants chosen to satisfy the sum condition. In particular, we should take

$$\lambda_j = \frac{\lambda}{2} 2^{kr}(1 - 2^{-r})2^{-jr}$$

for some $r < \frac{1}{p}$. If we do this, then for some λ large enough we will be able to assume that j is large enough to apply the Gaussian large deviations tail bound, which shows that the probability on the right-hand side in (5) is at most

$$\exp\left(-\Omega\left(\lambda_j^2 2^{j(1+\frac{2}{p})}\right)\right) \leq \exp\left(-\Omega\left(\lambda^2 2^{2kr} 2^{-2jr} 2^{j(1+\frac{2}{p})}\right)\right)$$

for some constant c. In particular, this is decaying faster than exponentially in terms of j, meaning that if we sum this over j then the entire sum will be dominated by the first term, i.e. we get

$$\mathbb{P}\left[\|u_{\geq k}\| > \frac{\lambda}{2}\right] \leq \exp\left(-\Omega\left(\lambda^2 2^{k(1+\frac{2}{p})}\right)\right).$$

Finally, plugging in our value of k from (4), we find that

$$\begin{aligned}\mathbb{P}[\|u\|_p > \lambda, \|u\|_2 \leq M] &\leq \exp\left(-\Omega\left(\lambda^2 \lambda^{\frac{2p}{p-2}(1+\frac{2}{p})} M^{-\frac{2p}{p-2}(1+\frac{2}{p})}\right)\right) \\ &= \exp\left(-\Omega\left(\lambda^{\frac{4p}{p-2}} M^{-2\frac{p+2}{p-2}}\right)\right).\end{aligned}$$

Recall that we wanted this probability to be $\leq Ce^{-c\lambda^p}$ for some $c > \frac{1}{p}$. If $p < 6$, then the power of λ is already < 6, meaning we are done. If $p = 6$, then the power of λ is equal to 6, but since M is taken to a negative power, we are good as long as M is small enough.