Quasi-Periodic Solutions to Nonlinear PDE's

W.-M. Wang

IAS, Oct 26, 2017

I. Introduction

In this talk, we shall consider nonlinear PDE's, such as the nonlinear Schrödinger (NLS):

$$i\frac{\partial u}{\partial t} + \Delta u + |u|^{2p}u + h.o.t. = 0;$$

and the nonlinear wave equations (NLW):

$$\frac{\partial^2 u}{\partial t^2} - \Delta u + u + u^{2p+1} + h.o.t. = 0$$

where $p \in \mathbb{N}$ is arbitrary.

Without the nonlinear terms, the linear Schrödinger equation

$$i\frac{\partial u}{\partial t} + \Delta u = 0;$$

describes the motion of a free particle. The nonlinear terms in NLS approximate the many body interactions. Similarly NLW (NLKG) describes the relativistic motion of a particle taking into account also the effect of many body interactions. They are among the most important and most basic equations coming from physics.

The dynamics of the linear Schrödinger and wave equations are completely described by the corresponding eigenfunctions (and eigenvalues) – the eigenfunctions form a spanning set and **all** solutions are their linear superpositions.

In the presence of the nonlinear term, there is no longer this linear superposition. Simplest generalizations of eigenfunctions are the time-periodic solutions (to the nonlinear equations). The next generalizations on the ladder are the quasi-periodic solutions, solutions which are periodic with several frequencies.

Example. When α is irrational, the function

$$f = \cos t + \cos \alpha t$$

is quasi-periodic in time with two frequencies.

Quasi-periodic solutions can be seen as bifurcations of linear superpositions of eigenfunctions. (If there is only one eigenfunction, then bifurcation leads to periodic solutions, as mentioned earlier.) They are hence quite natural. Clearly not all linear solutions will bifurcate to nonlinear solutions.

So one of our tasks is to impose reasonable conditions on the linear solutions, which are often algebro-geometric and arithmetic, as we shall explain, so that they could become solutions to the nonlinear equations, after small deformation. (This algebraic aspect could also be understood from a variational point of view, more on this later.)

Technical point: The algebro-geometric conditions tend to be on the eigenfunctions and the arithmetic ones on the eigenvalues. Typically the arithmetic conditions entail spacing of appropriate *linear combinations* of eigenvalues; while the algebro-geometric conditions are akin to geometric ray analysis.

Generally speaking, in order to have solutions which are quasi-periodic, the equations need to be posed on a compact manifold. The most basic is perhaps the flat torus $\mathbb{T}^d$, which leads to time quasi-periodic but space periodic solutions.

The plan of the talk is to give an idea of the proof of the following:

Theorem. There is a class of global solutions to the NLS and NLW in arbitrary dimension d and for arbitrary nonlinearity p. These are quasi-periodic solutions, which are Gevrey functions in space and time.

• Along the proof, we establish a **new** method, which is based on Anderson localization (multiscale analysis of Fröhlich and Spencer), harmonic analysis (Bourgain, Bourgain-Goldstein-Schlag) and algebraic analysis (W).

• These quasi-periodic solutions are also called KAM, Kolmogorov-Arnold-Moser, solutions (tori). KAM solutions are generally obtained using the Hamiltonian formalism, i.e., in the phase space; while our multiscale method fits naturally in the Lagrangian formalism, i.e. in the configuration space.

• For example, NLS is first-order in time and naturally in the Hamiltonian formalism; while NLW second-order and naturally Lagrangian. (Of course NLW also has a Hamiltonian representation as a system of first order equations with u and $\dot{u}$ as conjugate variables, but there is an analysis subtlety involved ...)

• Some historical contexts: After the success of KAM (Kolmogorov-Arnold-Moser) theory, Moser proposed to construct quasi-periodic solutions in the Lagrangian formalism. This is natural as the Lagrangian formulation could be more general. (For example, it requires less smoothness and deals more naturally with higher order equations.)

Based on Moser's notes, Salamon and Zehnder (ETH) showed the existence of quasi-periodic solutions using the Lagrangian formalism. The main point is to overcome small-divisors in order to obtain an implicit function theorem. This is in finite dimensions.

Our Theorem deals with PDE's, and is therefore in infinite dimensions. The proof uses Multiscale Analysis to control small-divisors.

II. The torus setting

Coming back to the problem at hand, we start from the linear equations posed on the torus $\mathbb{T}^d$:

$$i\frac{\partial u}{\partial t} + \Delta u = 0;$$

and

$$\frac{\partial^2 u}{\partial t^2} - \Delta u + u = 0,$$

where the term 1 is added, in order to avoid the singularity at 0 and is usually called the mass.

These linear equations are solved using Fourier series and spectral theory. The spectrum of the Laplacian:

$$\sigma(\Delta) = \{j_1^2 + j_2^2 + ... + j_d^2 \,|\, j_k \in \mathbb{Z}, k = 1, 2, ..., d\};$$

while that of the wave operator D, $D = \sqrt{-\Delta + 1}$:

$$\sigma(D) = \{\sqrt{j_1^2 + j_2^2 + ... + j_d^2 + 1} \,|\, j_k \in \mathbb{Z}, k = 1, 2, ..., d\}.$$

They play an essential role in our analysis. We note that in $d \geq 2$, they are highly **degenerate**. The spectrum of the wave operator D is moreover **dense**, which makes NLW even more difficult.

Below we explain our method using NLW, as not only it is more difficult but it also represents a generic case. (According to a theorem of Duistermaat and Guillemin, the spectrum of the differences of eigenvalues of a first order elliptic operator on a generic compact manifold is dense.) Since we seek real solutions, we use cosine series. (For NLS, complex solutions and Fourier series.)

III. The space-time cosine series

The linear equation admits solutions of the form

$$\cos(-(\sqrt{j^2+1})t + j\cdot x),\ j\in\mathbb{Z}^d.$$

More generally, let $u^{(0)}$ be a solution of finite number of frequencies, b frequencies, to the linear equation:

$$u^{(0)}(t,x) = \sum_{k=1}^{b} a_k \cos(-(\sqrt{j_k^2+1})t + j_k\cdot x).$$

Examples.

To illustrate the notations, let us fix: $d = 3$ and $b = 4$.
Then there are 4 vectors in $\mathbb{Z}^3$, e.g.,
$j_1 = (1, 2, 3)$, $j_2 = (5, 6, 7)$, $j_3 = (17, 41, 35)$ and $j_4 = (113, 27, 9)$,
which are the (vector) space frequencies;

and the 4 (scalar) time frequencies:
$\omega_1 = \sqrt{j_1^2 + 1} = \sqrt{15}$, $\omega_2 = \sqrt{j_2^2 + 1} = \sqrt{111}$,
$\omega_3 = \sqrt{j_3^2 + 1} = \sqrt{3196}$, $\omega_4 = \sqrt{j_4^2 + 1} = \sqrt{13580}$,
which are sometimes written in the vector notation:
$\omega = (\omega_1, \omega_2, \omega_3, \omega_4) \in \mathbb{R}^4$.

Note that at this stage, the frequencies $\omega := \omega^{(0)}$ are **fixed**, because $u^{(0)}$ is a solution to the linear equation.

$$\frac{\partial^2 u}{\partial t^2} - \Delta u + u = 0.$$

Time quasi-periodic solutions, maybe viewed as "periodic on higher dimensional torus": to each frequency, we add a dimension; so as an ansatz, we may seek solutions to the nonlinear equations in the form:

$$u(t,x) = \sum_{(n,j)\in\mathbb{Z}^{b+d}} \hat{u}(n,j)\cos(n\cdot\omega t + j\cdot x)$$

where b is the number of basic frequencies in time, with the frequency $\omega \in \mathbb{R}^b$ to be determined.

In this formulation, $u^{(0)}$ may be written in the form:

$$u^{(0)}(t,x) = \sum_{k=1}^{b} a_k \cos(-(\sqrt{j_k^2+1})t + j_k \cdot x)$$
$$:= \sum_{k=1}^{b} \hat{u}^{(0)}(\mp e_k, \pm j_k) \cos(\mp e_k \cdot \omega^{(0)} t \pm j_k \cdot x),$$

where $e_k = (0, 0, ...1, .., 0) \in \mathbb{Z}^b$ is a unit vector, with the only non-zero component in the kth direction, $\omega_k^{(0)} = \sqrt{j_k^2+1}$ and

$$\hat{u}(-e_k, j_k) = \hat{u}(e_k, -j_k) = a_k/2.$$

Remark. The time part of the Fourier series can be understood using the substitution:

$$\partial/\partial t \mapsto \sum_{i=1}^{b} \omega_i \partial/\partial y_i,$$

where y_i are the variables on the higher dimensional torus in time.

IV. The matrix problem and the Anderson model

Using the cosine series ansatz, the linear Klein-Gordon operator:

$$\frac{\partial^2 u}{\partial t^2} - \Delta u + u,$$

becomes the diagonal matrix operator

$$\text{diag}\,[(n \cdot \omega)^2 - j^2 - 1],$$

on $\ell^2(\mathbb{Z}^{b+d})$.

As mentioned earlier, ω is to be determined. Since we perturb from the linear solution $u^{(0)}$. As an initial approximation $\omega = \omega^{(0)}$ are the linear frequencies $\sqrt{j_k^2 + 1}$, $k = 1, 2, ..., b$.

So the diagonals can be 0, e.g., when $|n| = 1$ and $j = j_k$, $k = 1, 2, ..., b$. The off-diagonal matrix is a convolution matrix, since the nonlinear term u^{2p+1} is multiplicative.

Since we seek small solutions, we may use a Newton scheme and linearize about $u^{(0)}$, leading to the linear self-adjoint matrix operator:

$$T = D + P = \text{diag}[(n \cdot \omega^{(0)})^2 - j^2 - 1] + \hat{u}^{(0)} * \hat{u}^{(0)},$$

where for illustration, we have set $p = 1$.

To illustrate the convolution matrix P, let us take a simplified example and assume $u := u^{(0)} = \cos x$, $x \in \mathbb{R}$. Then since

$$u^2 = \cos^2 x = \frac{\cos 2x - 1}{2},$$

P has diagonal entries $-1/2$ and two off-off diagonal entries $1/2$.

The Anderson models

The Anderson models are, on the other hand, as follows:

(a) The random Anderson operator:

$$H = V + \Delta = \text{diag}\, v_i + \Delta,$$

where $\{v_i\}$ are a family of iid random variables and Δ is the discrete Laplacian, as in the 1983 foundational paper of Fröhlich and Spencer; and

(b) The quasi-periodic Anderson operator:

$$H = V(n\omega) + \Delta,$$

where ω is irrational. For example, in 1d,

$$V = \cos(n\omega + \theta),$$

which leads to the celebrated almost Mathieu operator; and in 2d, a special case of Anderson localization studied in the seminal paper of Bourgain-Goldstein-Schlag.

Now the $H's$ and the T look alike, particularly the quasi-periodic H, except ω is a parameter in H; while $\omega = \omega^{(0)}$ is **fixed** in T.

So let us present first the Fröhlich-Spencer Anderson localization scheme for parameter dependent H, and then explain the additional (algebro) geometric step needed to deal with ω, which is, to leading order, fixed.

The scheme invented in [FS] is a multiscale analysis, using appropriate finite sets Λ to approximate $\mathbb{Z}^d$, and consists of 3 steps:

1. Initial scale (or scales) Green's function estimates in a (sufficiently large) box;

2. Cover larger box by smaller boxes and obtain Green's function estimates for the larger box. Central is that there are only, but "a few" bad small boxes, where the Green's function estimates are not good, but still there exist a priori estimates (Wegner estimates);

3. Use the Green's function estimates to prove that the spectrum of H is pure point ..., i.e., AL.

Returning to the problem at hand:

V. The concept of the bi-characteristics

In the nonlinear construction to prove the Theorem, the key is the invertibility of the (truncated) linearized operators, i.e., the Green's function estimates.

We call the scales in Step 2, the asymptotic regions. The iterative scheme in the asymptotic regions are the same for T and H. This part is universal.

However, the initial estimates (Step 1) are **different**. This is because in T, the diagonals can be 0, since there are no parameters at this stage; while using the parameters ω, the diagonals in H can be kept away from 0 by excise a **small** set in ω (assuming V is large).

In other words, the perturbation theory for T is singular.

For example, for the parameter dependent quasi-periodic Anderson model, a singular case would be the model:

$$H = V(\delta n\omega) + \delta\Delta,$$

with δ small, at some initial scale.

Generally speaking, in the singular case, the estimates need to be more quantitative.

In our case, analysis alone does not suffice to deal with singular perturbation, we need to analyze the geometry of the zero-set:

$$\mathcal{C} := \{(n,j) \in \mathbb{Z}^{b+d} | (n\cdot\omega)^2 - j^2 - 1 = 0\}.$$

We call the above set, the bi-characteristics. Its geometry will play an essential role in estimating T^{-1}. (This step is **non-perturbative**.)

VI. A separation property on the bi-characteristics

For H, the "bi-characteristics" set $\mathcal{C} = \emptyset$, after excision in ω. This is not possible for T. The next best thing is then trying to achieve some separation property (clustering property) on $\mathcal{C}$.

The key question becomes: separated, but relative to what ??

Answer: relative to a precise notion of connected, which we define below.

Recall that

$$T = D + P$$

and P is a convolution matrix. So there is a natural notion of connected, namely, we say two points x and $y \in \mathbb{Z}^{b+d}$ are connected if

$$(x - y) \in \text{supp } P.$$

We want then that the connected sets on the bi-characteristics $\mathcal{C}$ are small. Now due to the nature of singular perturbation theory, the control on the size needs to be quantitatively optimal.

We obtain this quantitative control by making an algebraic description of the connected sets and variable reductions as follows. (Below we write ω for $\omega^{(0)}$.)

• The algebraic description: by definition, if $(n, j) \in \mathcal{C}$ and $(n', j') \in \mathcal{C}$ are connected, then the following two equations are satisfied:

$$(n \cdot \omega)^2 - j^2 - 1 = 0,$$

$$[(n + \nu) \cdot \omega]^2 - (j + \eta)^2 - 1 = 0,$$

for some $(\nu, \eta) \in \text{supp } P$.

In fact, it is easier to illustrate the idea on Schrödinger, so let us first do that. The characteristics for Schrödinger is:

$$\mathcal{C} := \{(n, j) \in \mathbb{Z}^{b+d} | n \cdot \omega + j^2 = 0\}.$$

So if $(n, j) \in \mathcal{C}$ and $(n', j') \in \mathcal{C}$ are connected, then the following two equations are satisfied:

$$n \cdot \omega + j^2 = 0,$$

$$(n + \nu) \cdot \omega + (j + \eta)^2 = 0,$$

for some $(\nu, \eta) \in \text{supp } P$.

Subtracting the two equations leads to a linear equation in j, i.e., a hyperplane, parametrized by (ν, η), which are determined by $u^{(0)}$, the solution to the linear equation.

So a block, a connected set can be described by a system of linear equations; the size of a block, the size of a connecte set is then bounded above by the size of the corresponding compatible linear system.

Under appropriate geometric conditions on $u^{(0)}$, variable reductions give that the size is at most $2d$, which is essentially optimal. (Recall that j is a d-vector and there are two characteristics.)

For Wave, the same reasoning leads to hyperboloids. The variable reductions are much more difficult, because the hyperboloids

$$(n \cdot \omega)^2 - j^2 - 1 = 0,$$

are essentially **flat** away from the origin, compatible with convolution (translation invariance). So aside from geometric conditions akin to that for Schrödinger, arithmetic conditions also come into play.

VIII. The Diophantine properties of ω

Recall that the frequencies

$$\omega_k = \sqrt{j_k^2 + 1},$$

where

$$j_k^2 := j_{k,1}^2 + j_{k,2}^2 + ... + j_{k,d}^2.$$

• We impose the condition that $(j_k^2 + 1)$ $(j_k \neq 0)$ are distinct and **square-free**, for $k = 1, 2, ..., b$.

As a consequence, there is the usual (linear) independence:

$$\|n \cdot \omega\|_{\mathbb{T}} \neq 0,$$

where $n \neq 0$;

as well as the **quadratic** non-equality:

$$\|\sum_{k<\ell} n_k n_\ell \omega_k \omega_\ell\|_{\mathbb{T}} \neq 0,$$

where $\sum_{k<\ell} |n_k n_\ell| \neq 0$.

The linear independence is familiar – it takes care of the (usual) small-divisors arising from multi-scale analysis.

The quadratic non-equality is new and takes care of hyperbolicity; moreover it **doubles** as a small-divisor lower bound for the dense linear flow:

If

$$n \cdot \omega + \sqrt{j^2 + 1} \neq 0,$$

then

$$|n \cdot \omega + \sqrt{j^2 + 1}| > c|n|^{-\alpha}, \ |n| \neq 0,$$

where c, $\alpha > 0$, cf. [W. Schmidt].

It is worth noting that these are new types of small-divisors, which are geometric in origin and do not appear in NLS. The **quadratic** Diophantine property arises because NLW is a **second** order in time equation. This appears to be new.

Once the separation is achieved, one can **extract parameters** from the nonlinearity – the parameters are the amplitudes a, the Fourier coefficients of the linear solution $u^{(0)}$. Bourgain's non-resonant method becomes available to deal with small-divisors.

There is, however, a new analysis point in our application, namely for measure estimates, it is more convenient to lower regularity and work in the Lipschitz setting. (Recall that the Theorem is in C^∞-Gevrey setting.) With this addition, one may then proceed to prove the Theorem.

It should be mentioned that after the geometric step of extracting parameters, the ideas for the linear analysis is modelled after that of quasi-periodic AL of Bourgain-Goldstein-Schlag.

IX. Literature

• NLS with external parameters ω: Bourgain (AL-Lagrangian), 1998; Eliasson-Kuksin (KAM-normal form method = infinite sequence of change of coordinates), 2010

• NLS: Procesi-Procesi (KAM method), 2016; Wang (AL-Lagrangian), 2016

• NLW with external parameters ω: Bourgain (sketch of proof), 2005

• NLW: Wang (AL-Lagrangian), 2017