Dedicated to Walter Craig on his 60-th birthday

TIME-AVERAGING FOR WEAKLY NONLINEAR CGL EQUATIONS WITH ARBITRARY POTENTIALS

HUANG GUAN, KUKSIN SERGEI, AND MAIOCCHI ALBERTO

ABSTRACT. Consider weakly nonlinear complex Ginzburg–Landau (CGL) equation of the form:

$$u_t + i(-\triangle u + V(x)u) = \epsilon\mu\Delta u + \epsilon\mathcal{P}(\nabla u, u), \quad x \in \mathbb{R}^d\,, \qquad (*)$$

under the periodic boundary conditions, where $\mu \geqslant 0$ and $\mathcal{P}$ is a smooth function. Let $\{\zeta_1(x), \zeta_2(x), \dots\}$ be the L_2-basis formed by eigenfunctions of the operator $-\triangle + V(x)$. For a complex function $u(x)$, write it as $u(x) = \sum_{k\geqslant 1} v_k\zeta_k(x)$ and set $I_k(u) = \frac{1}{2}|v_k|^2$. Then for any solution $u(t,x)$ of the linear equation $(*)_{\epsilon=0}$ we have $I(u(t,\cdot)) = const$. In this work it is proved that if equation $(*)$ with a sufficiently smooth real potential $V(x)$ is well posed on time-intervals $t \lesssim \epsilon^{-1}$, then for any its solution $u^\epsilon(t,x)$, the limiting behavior of the curve $I(u^\epsilon(t,\cdot))$ on time intervals of order ϵ^{-1}, as $\epsilon \to 0$, can be uniquely characterized by a solution of a certain well-posed effective equation:

$$u_t = \epsilon\mu\triangle u + \epsilon F(u),$$

where $F(u)$ is a resonant averaging of the nonlinearity $\mathcal{P}(\nabla u, u)$. We also prove similar results for the stochastically perturbed equation, when a white in time and smooth in x random force of order $\sqrt{\epsilon}$ is added to the right-hand side of the equation.

The approach of this work is rather general. In particular, it applies to equations in bounded domains in $\mathbb{R}^d$ under Dirichlet boundary conditions.

1. INTRODUCTION

Equations. We consider a weakly nonlinear CGL equation on a rectangular d-torus $T^d = \mathbb{R}/(L_1\mathbb{Z}) \times \mathbb{R}/(L_2\mathbb{Z}) \times \cdots \times \mathbb{R}/(L_d\mathbb{Z})$, $L_1, \dots, L_d > 0$,

$$u_t + i(-\Delta + V(x))u = \epsilon\mu\Delta u + \epsilon\mathcal{P}(\nabla u, u), \quad u = u(t,x),\ x \in T^d, \qquad (1.1)$$

where $\mu \geqslant 0$, $\mathcal{P} : \mathbb{C}^{d+1} \to \mathbb{C}$ is a C^∞-smooth function, ϵ is a small parameter and $V(\cdot) \in C^n(T^d)$ is a sufficiently smooth real-valued function on T^d (we will assume that n is large enough). If $\mu = 0$, then the nonlinearity $\mathcal{P}$ should be independent of the derivatives of the unknown function u. For simplicity, we assume that $\mu > 0$. The case $\mu = 0$ can be treated exactly in the same way (even simpler).

For any $s \in \mathbb{R}$ we denote by H^s the Sobolev space of complex-valued functions on T^d, provided with the norm $\|\cdot\|_s$,

$$\|u\|_s^2 = \langle(-\Delta)^s u, u\rangle + \langle u, u\rangle, \quad \text{if} \quad s \geq 0\,,$$

where $\langle\cdot,\cdot\rangle$ is the real scalar product in $L^2(T^d)$,

$$\langle u, v\rangle = \mathrm{Re}\int_{T^d} u\bar{v}\mathrm{d}x, \quad u, v \in L^2(T^d).$$

For any $s > d/2+1$, it is known that the mapping $\mathcal{P} : H^s \to H^{s-1},\ u \mapsto \mathcal{P}(\nabla u, u)$, is smooth and locally Lipschitz, see below Lemma 3.1.

Our goal is to study the dynamics of Eq. (1.1) on time intervals of order ϵ^{-1} when $0 < \epsilon \ll 1$. Introducing the slow time $\tau = \epsilon t$, we rewrite the equation as

$$\dot{u} + \epsilon^{-1} i(-\Delta + V(x))u = \mu\Delta u + \mathcal{P}(\nabla u, u), \tag{1.2}$$

where $u = u(\tau, x)$, $x \in T^d$, and the upper dot $\dot{}$ stands for $\frac{\mathrm{d}}{\mathrm{d}\tau}$. We assume

Assumption A: *There exists a number $s_* \in (d/2+1, n]$ and for every $M_0 > 0$ there exists $T = T(s_*, M_0) > 0$ such that if $u_0 \in H^{s_*}$ and $\|u_0\|_{s_*} \leq M_0$, then Eq. (1.2) has a unique solution $u(\tau,x) \in C([0,T], H^{s_*})$ with the initial datum u_0, and $||u(\tau,x)||_{s_*} \leqslant C(s_*, M_0, T)$ for $\tau \in [0,T]$.*

This assumption can be verified for Eq. (1.1) with various nonlinearities $\mathcal{P}$. For example when $\mu = 0$ and $V(x) \equiv 0$ it holds if $\mathcal{P}(u)$ is any smooth function. Indeed, taking the scalar product in the space H^{s_*} of eq. (1.2) with $u(t)$ and using the Granwall lemma we get the Assumption A with suitable positive constants $T(s_*, M_0)$ and $C(s_*, M_0, T)$. When $\mu > 0$, the assumption with any $T > 0$ is satisfied by Eq. (1.1) with nonlinearity $\mathcal{P}(u) = -\gamma_R f_p(|u|^2)u - i\gamma_I f_q(|u|^2)u$, where $\gamma_R, \gamma_I > 0$, the functions $f_p(r)$ and $f_q(r)$ are the monomials $|r|^p$ and $|r|^q$, smoothed out near zero, and

$$0 \leqslant p, q < \infty \quad \text{if} \quad d = 1,2 \quad \text{and} \quad 0 \leqslant p, q < \min\left\{\frac{d}{2}, \frac{2}{d-2}\right\} \quad \text{if} \quad d \geqslant 3,$$

see, e.g. [8].

We denote by A_V the Schrödinger operator

$$A_V u := -\Delta u + V(x)u.$$

Let $\{\lambda_k\}_{k\geqslant 1}$ be its eigenvalues, ordered in such a way that

$$\lambda_1 \leqslant \lambda_2 \leqslant \lambda_3 \leqslant \cdots,$$

and let $\{\zeta_k,\ k \geqslant 1\}$ of $L^2(T^d)$ be an orthonormal basis, formed by the corresponding eigenfunctions. We denote $\Lambda = (\lambda_1, \lambda_2, \dots)$ and call Λ the *frequency vector* of Eq. (1.2). For a complex-valued function $u \in H^s$, we denote by

$$\Psi(u) := v = (v_1, v_2, \dots), \quad v_k \in \mathbb{C}, \tag{1.3}$$

the vector of its Fourier coefficients with respect to the basis $\{\zeta_k\}_{k\geqslant 1}$: $u = \sum_{k\geqslant 1} v_k \zeta_k$. Note that Ψ is a real operator: it maps real functions $u(x)$ to real vectors v. In the space of complex sequences $v = (v_1, v_2, \dots)$, we introduce the norms

$$|v|_s^2 = \sum_{k=1}^{+\infty} (|\lambda_k|^s + 1)\, |v_k|^2, \quad s \in \mathbb{R},$$

and denote $h^s = \{v : |v|_s < \infty\}$. Clearly Ψ defines an isomorphism between the spaces H^s and h^s.

Now we write Eq. (1.2) in the v-variables:

$$\dot{v}_k + \epsilon^{-1} i\lambda_k v_k = -\mu\lambda_k v_k + P_k(v), \quad k \in \mathbb{N}, \tag{1.4}$$

where

$$P(v) := (P_k(v),\ k \in \mathbb{N}) = \Psi\Big(\mu V(x)u + \mathcal{P}(\nabla u, u)\Big), \quad u = \Psi^{-1} v. \tag{1.5}$$

For every $k \in \mathbb{N}$ we set

$$I_k(v) = \frac{1}{2} v_k \bar{v}_k, \quad \text{and } \varphi_k(v) = \text{Arg } v_k \in \mathbb{T}^1 = \mathbb{R}/(2\pi\mathbb{Z}) \text{ if } v_k \neq 0, \text{ else } \varphi_k = 0. \tag{1.6}$$

Then $v_k = \sqrt{2I_k} e^{i\varphi_k}$. Notice that the quantities I_k are conservation laws of the linear equation $(1.1)_{\epsilon=0}$, and that the variables $(I, \varphi) \in \mathbb{R}_+^\infty \times \mathbb{T}^\infty$ are its action-angles. For any $(I, \varphi) \in \mathbb{R}_+^\infty \times \mathbb{T}^\infty$ we denote

$$v = v(I, \varphi) \quad \text{if} \quad v_k = \sqrt{2I_k} e^{i\varphi_k}, \ \ \forall\, k\,. \tag{1.7}$$

If this relation holds, we will write $v \sim (I, \varphi)$. We introduce the weighted l^1-space h_I^s:

$$h_I^s := \{I = (I_k,\ k \in \mathbb{N}) \in \mathbb{R}^\infty : |I|_s^\sim = \sum_{k=1}^{+\infty} 2(|\lambda_k|^s + 1)|I_k| < \infty\}.$$

Then $|v|_s^2 = |I(v)|_s^\sim$, for each $v \in h^s$. Using the action-angle variables (I, φ), we write Eq. (1.4) as a slow-fast system:

$$\dot{I}_k = v_k \cdot \big(-\mu\lambda_k v_k + P_k(v) \big), \quad \dot{\varphi}_k = -\epsilon^{-1}\lambda_k + |v_k|^{-2} \cdots, \quad k \in \mathbb{N}.$$

Here $a \cdot b$ denotes $\text{Re}(a\bar{b})$, for $a, b \in \mathbb{C}$, and the dots stand for a factor of order 1 (as $\epsilon \to 0$).

Effective equations. Our task is to study the evolution of the actions I_k when $\epsilon \ll 1$ and $0 \leq \tau \lesssim 1$. An efficient way to deal with this problem is through the so-called *interaction representation.* Let us define

$$a_k(\tau) = e^{i\epsilon^{-1}\lambda_k \tau} v_k(\tau)\ . \tag{1.8}$$

Then

$$|a_k|^2 = |v_k|^2 = 2I_k\ , \tag{1.9}$$

so to study the evolution of the actions we can use the a-variables instead of the v-variables. Using Eq. (1.4), we obtain for $a = (a_1, a_2, \dots)$ the system of equations

$$\dot{a}_k(\tau) = -\mu\lambda_k a_k + e^{i\epsilon^{-1}\lambda_k\tau} P_k(\Phi_{-\epsilon^{-1}\Lambda\tau} a), \quad k \in \mathbb{N}\,, \tag{1.10}$$

where for each $\theta = (\theta_k,\ k \in \mathbb{N}) \in \mathbb{R}^\infty$, Φ_θ stands for the linear operator in h^s defined by

$$\Phi_\theta v = v', \quad v'_k = e^{i\theta_k} v_k \quad \forall\, k\,.$$

Clearly Φ_θ defines isometries of all Hilbert spaces h^s, and in the action-angle variables it reads $\Phi_\theta(I, \varphi) = (I, \varphi + \theta)$.

To approximately describe the dynamics of Eq. (1.10) with $\epsilon \ll 1$ we introduce an effective equation:

$$\dot{\tilde{a}}_k = -\mu\lambda_k \tilde{a}_k + R_k(\tilde{a}), \quad k \in \mathbb{N}, \tag{1.11}$$

where $R(\tilde{a}) := (R_k(\tilde{a}),\ k \in \mathbb{N})$ and

$$R(\tilde{a}) = \lim_{T \to \infty} \frac{1}{T} \int_0^T \Phi_{\Lambda t} P(\Phi_{-\Lambda t} \tilde{a}) dt. \tag{1.12}$$

We will see in Sections 2 and 3 that the limit in (1.12) is well defined and that Eq. (1.11) is well posed, at least locally in time.

Results. In Section 4 we prove that the actions of solutions for the effective equation approximate well the actions $I_k(v(\tau))$ of solutions v for (1.4). Let us fix any $M_0 > 0$.

Theorem 1.1. *Let $u(\tau,x)$, $0\le\tau\le T=T(s_*,M_0)$, be a solution of* (1.2)*, such that $u(0,x)=u_0(x)$, $\|u_0\|_{s_*}\le M_0$, existing by Assumption A. Denote $v(\tau)=\Psi(u(\tau,\cdot))$, $0\le\tau\le T$. Then a solution $\tilde a(\tau)$ of* (1.11)*, such that $\tilde a(0)=v(0)$, exists for $0\le\tau\le T$, and for any $s_1<s_*$ we have*

$$\sup_{0\le\tau\le T}|I(v(\tau))-I(\tilde a(\tau))|^{\sim}_{s_1}\to 0,\quad \textit{as}\quad \epsilon\to 0\,.$$

The rate of the convergence does not depend on u_0, if $\|u_0\|_{s_*}\le M_0$.

This theorem may be regarded as a PDE-version of the Bogolyubov averaging principle, see [3] and [1], Section 6.1. The result and its its proof may be easily recasted to a theorem on perturbations of linear Hamiltonian systems with discrete spectrum. Instead of doing this, below we briefly discuss its generalisations to other nonlinear PDE problems.

In the second part of the paper (Sections 5–7) we consider the CGL equations (1.1) with added small random force:

$$u_t+i(-\Delta+V(x))u=\epsilon\mu\Delta u+\epsilon\mathcal{P}(\nabla u,u)+\sqrt{\epsilon}\,\frac{d}{dt}\sum_{l\ge1}b_l\beta_l(t)e_l(x),\tag{1.13}$$

where $u=u(t,x)$, $x\in T^d$, the coefficients b_l decay fast enough with $|l|$, $\{\beta_l(t)\}$ are standard independent complex Wiener processes and $\{e_l(x)\}$ is the usual trigonometric basis of the space $L_2(T^d)$, parametrized by natural numbers. It turns out that the effective equation for (1.13) is the equation (1.11), perturbed by a suitable stochastic forcing, see Section 5. Assuming that the function $\mathcal{P}$ has at most a polynomial growth and that the equation satisfies a suitable stochastic analogy of the Assumption A we prove a natural stochastic version of Theorem 1.1 (see Theorem 5.2). Next, supposing that the stochastic effective equation is mixing and has a unique stationary measure μ_0, we prove in Theorem 5.4 that if μ_ϵ is a stationary measure for Eq. (1.13), then $\Psi\circ\mu_\epsilon$ converge to μ_0 as $\epsilon\to0$. So if the stochastic effective equation is mixing, then it comprises asymptotical properties of solutions for Eq. (1.13) as $t\to\infty$ and $\epsilon\to0$.

The proof of the theorems in this work follows the Anosov approach to averaging in finite-dimensional systems (see in [1, 18]), its version for averaging in resonant systems (see in [1]) and its stochastic version due to Khasminski [12]. The crucial idea that for averaging in PDEs the averaged equations for actions (which are equations with singularities) should be considered jointly with suitable effective equations (which are regular equations) was suggested in [13] for averaging in stochastic PDEs, and later was used in [14] and [8, 9, 15, 16]. It was realised in the second group of publications that for perturbations of linear systems the method may be well combined with the interaction representation of solutions, well known and popular in nonlinear physics (see [3, 19]), and which already was used for purposes of completely resonant averaging, corresponding to constant coefficient PDEs with small nonlinearities on the square torus (see [7, 5]).

For the case when the spectrum of the unperturbed linear system is non-resonant (see below Example 2.2), the results of this paper were obtained in [14, 8], while for the case when the spectrum is completely resonant – in [15, 9]. The novelty of this work is a version of the Anosov method of averaging, applicable to nonlinear PDEs with small nonlinearities, which does not impose restrictions on the spectrum of the unperturbed equation.

Alternatively, the averaging for weakly nonlinear PDEs may be studied, using the normal form techniques, e.g. see [2] and references therein. Compared to the Anosov approach, exploited in this work, the method of normal form is much more demanding to the spectrum of the unperturbed equation, and more sensitive to its perturbations. So usually it applies only in small vicinities of equilibriums. Its advantage is that it may imply stability on longer time intervals, while the method of this work is restricted to the first-order averaging. So in the deterministic setting it allows to control solutions of ϵ-perturbed equations only on time-intervals of order ϵ^{-1} (still, in the stochastic setting it also allows to control the stationary measure, which describes the asymptotic behaviour of solutions as $t \to \infty$).

Generalizations. The Anosov-like method of resonant averaging, presented in this work, is very flexible. With some slight changes, it easily generalizes to weakly nonlinear CGL equations, involving high order derivatives,

$$u_t + i(-\triangle u + V(x)u) = \epsilon \mathcal{P}(\nabla^2 u, \nabla u, u, x), \quad x \in T^d, \tag{1.14}$$

provided that the Assumption A holds and the corresponding effective equation is well posed locally in time. See in Appendix A (also see [8], where a similar result is proven for the case of non-resonant spectra).

The method applies to equations (1.1) and (1.13) in a bounded domain $\mathcal{O} \subset \mathbb{R}^d$ under Dirichlet boundary conditions. Indeed, if $d \leq 3$, then to treat the corresponding boundary-value problem we can literally repeat the argument of this work, replacing there the space H^s with the Hilbert space $H_0^2(\mathcal{O}) = \{u \in H^2(\mathcal{O}) : u \mid_{\partial\mathcal{O}} = 0\}$. If $d \geq 4$, then H^s should be replaced with an L_p-based Banach space $W_0^{2,p}(\mathcal{O})$, where $p > d/2$.

Obviously the method applies to weakly nonlinear equations of other types; e.g. to weakly nonlinear wave equations. In [16] the method in its stochastic form was applied to the Hasegawa-Mima equation, regarded as a perturbation of the Rossby equation $(-\Delta + K)\psi_t(t, x, y) - \psi_x = 0$, while in [4] it is applied to systems of non-equilibrium statistical physics, where each particle is perturbed by an ε-small Langevin thermostat, and is studied the limit $\varepsilon \to 0$ (similar to the same limit in Eq. (1.13)).

The averaging for perturbations of nonlinear integrable PDEs is more complicated. Due to the lack in the functional phase-spaces of an analogy of the Lebesgue measure (required by the Anosov approach to the finite-dimensional deterministic averaging), in this case the results for stochastic perturbations are significantly stronger than the deterministic results. See in [10].

Acknowledgments. We are thankful to Anatoli Neishtadt for discussing the finite-dimensional averaging. This work was supported by l'Agence Nationale de la Recherche through the grant STOSYMAP (ANR 2011BS0101501).

2. Resonant averaging in Hilbert Spaces

The goal of this section is to show that the limit in (1.12) is well-defined in some suitable settings and study its properties. Below for an infinite-vector $v = (v_1, v_2, \dots)$ and any $m \in \mathbb{N}$ we denote

$$v^m = (v_1, \dots, v_m), \quad \text{or} \quad v^m = (v_1, \dots, v_m, 0, \dots),$$

depending on the context. This agreement also applies to elements $\varphi = (\varphi_1, \varphi_2, \dots)$ of the torus $\mathbb{T}^\infty$. For m-vectors I^m, φ^m, v^m we write $v^m \sim (I^m, \varphi^m)$ if (1.7) holds for $k = 1, \dots, m$. By Π^m, $m \geqslant 1$, we denote the Galerkin projection

$$\Pi^m : h^0 \to h^0, (v_1, v_2, \dots) \mapsto v^m = (v_1, \dots, v_m, 0, \dots).$$

For a continuous complex function f on a Hilbert space H, we say that f is locally Lipschitz and write $f \in Lip_{loc}(H)$ if

$$|f(v) - f(v')| \leqslant \mathcal{C}(R)\|v - v'\|, \quad \text{if} \quad \|v\|, \|v'\| \leqslant R, \tag{2.1}$$

for some continuous non-decreasing function $\mathcal{C} : \mathbb{R}^+ \to \mathbb{R}^+$ which depends on f. We write

$$f \in Lip_{\mathcal{C}}(H) \text{ if (2.1) holds and } |f(v)| \le \mathcal{C}(R) \text{ if } \|v\| \le R. \tag{2.2}$$

If $f \in Lip_{\mathcal{C}}(H)$, where $\mathcal{C}(\cdot) =$Const, then f is a bounded (globally) Lipschitz function. If B is a Banach space, then the space $Lip_{loc}(H, B)$ of locally Lipschitz mappings $H \to B$ and its subsets $Lip_{\mathcal{C}}(H, B)$ are defined similarly.

For any vector $W = (w_1, w_2, \dots) \in \mathbb{R}^\infty$ we set

$$\langle f \rangle^T_{W,l}(v) = \frac{1}{T}\int_0^T e^{iw_l t} f(\Phi_{-Wt}v)dt, \tag{2.3}$$

and if the limit of $\langle f \rangle^T_{W,l}(v)$ when $T \to \infty$ exists, we denote

$$\langle f \rangle_{W,l} = \lim_{T\to\infty} \langle f \rangle^T_{W,l}(v).$$

Concerning this definition we have the following lemma. Denote

$$B(M, h^s) = \{v \in h^s : |v|_s \leqslant M\}, \quad M > 0.$$

Lemma 2.1. *Let $f \in Lip_{\mathcal{C}}(h^{s_0})$ for some $s_0 \geqslant 0$ and some function $\mathcal{C}$ as above. Then*

(i) For every $T \neq 0$, $\langle f \rangle^T_{W,l} \in Lip_{\mathcal{C}}(h^{s_0})$.

(ii) The limit $\langle f \rangle_{W,l}(v)$ exists for $v \in h^{s_0}$ and this function also belongs to $Lip_{\mathcal{C}}(h^{s_0})$.

(iii) For $s > s_0$ and any $M > 0$, the functions $\langle f \rangle^T_{W,l}(v)$ converge, as $T \to \infty$, to $\langle f \rangle_{W,l}(v)$ uniformly for $v \in B(M, h^s)$.

(iv) The convergence is uniform for $f \in Lip_{\mathcal{C}}(h^{s_0})$ with a fixed function $\mathcal{C}$.

Proof. (i) It is obvious since the transformations Φ_θ are isometries of h^{s_0}.

(ii) To prove this, consider the restriction of f to $B(M, h^{s_0})$, for any fixed $M > 0$. Let us take some $v \in B(M, h^{s_0})$ and fix any $\rho > 0$. Below in this proof by $O(v)$, $O_1(v)$, etc, we denote various functions $g(v) = g(I, \varphi)$, defined for $|v|_{s_0} \le M$ and bounded by 1.

Let us choose any $m = m(\rho, M, v, \mathcal{C})$ such that

$$\mathcal{C}(M)\, |v - \Pi^m v|_{s_0} \le \rho\,.$$

Then $|f(v) - f(\Pi^m v)| < \rho$, and by (i)

$$\left|\langle f \rangle^T_{W,l}(v) - \langle f \rangle^T_{W,l}(\Pi^m v)\right| < \rho\,,$$

for every $T > 0$.

Let us set

$$\mathcal{F}^m(I^m, \varphi^m) = \mathcal{F}^m(v^m) = f\big(v^m\big), \qquad \forall\, v^m \sim (I^m, \varphi^m) \in \mathbb{C}^m\,,$$

where in the r.h.s. v^m is regarded as the vector $(v^m, 0, \dots)$. Clearly, the function $\varphi^m \mapsto \mathcal{F}^m(I^m, \varphi^m)$ is Lipschitz-continuous on $\mathbb{T}^m$. So its Fejer polynomials

$$\sigma_K(\mathcal{F}^m) = \sum_{k\in\mathbb{Z}^m,\, |k|_\infty \le K} a_k^K e^{ik\cdot(\varphi^m)}, \qquad K \ge 1\,,$$

where $a_k^K = a_k^K(m, I^m)$, converges to $\mathcal{F}^m(I^m, \varphi^m)$ uniformly on $\mathbb{T}^m$. Moreover, the rate of convergence depends only on its Lipschitzian norm and the dimension m (see e.g. Theorem 1.20, Chapter XVII of [22]). Therefore, there exists $K = K(\mathcal{C}, M, \rho, m) > 0$ such that

$$\mathcal{F}^m(I^m, \varphi^m) = \sum_{k\in\mathbb{Z}^m, |k|_\infty \leqslant K} a_k^K e^{ik\cdot\varphi^m} + \rho O_1(I^m, \varphi^m)\,. \tag{2.4}$$

Now we define

$$\mathcal{F}_K^{res}(I^m, \varphi^m) = \sum_{k\in S(K)} a_k^K e^{ik\cdot\varphi^m}, \quad S(K) = \{k \in \mathbb{Z}^m :\ |k|_\infty \leqslant K, w_l - \sum_{j=1}^m k_i w_i = 0\}.$$

Since

$$\mathcal{F}^m\big(\Phi_{-W^m t}(\Pi^m v)\big) = \mathcal{F}^m(I^m, \varphi^m - Wt)\,,$$

then

$$\langle e^{ik\cdot\varphi^m}\rangle_{W,l}^T = e^{ik\cdot\varphi^m} \quad \text{if} \quad k \in S(K)\,,$$

$$\Big|\langle e^{ik\cdot\varphi^m}\rangle_{W,l}^T\Big| \le \frac{2T^{-1}}{|w_l - k\cdot W^m|} \quad \text{if} \quad |k|_\infty \le K,\ k \notin S(K)\,,$$

where we regard $e^{ik\cdot\varphi^m}$ as a function of v. Accordingly,

$$\begin{aligned}\langle f\rangle_{W,l}^T(v) &= \langle \mathcal{F}^m(I^m, \varphi^m)\rangle_{W,l}^T + \rho O_2(v)\\ &= \mathcal{F}_K^{res}(I^m, \varphi^m) + C(\rho, M, W, f, I)T^{-1}O_3(v) + \rho O_4(v)\,.\end{aligned}$$

So there exists $\bar T = T(\rho, M, W, f, I) > 0$ such that if $T \geqslant \bar T$, then

$$\big|\langle f\rangle_{W,l}^T - \mathcal{F}_K^{res}(I^m, \varphi^m)\big| < 2\rho\,,$$

and for any $T' \geqslant T'' \geqslant \bar T$, we have

$$\big|\langle f\rangle_{W,l}^{T'}(v) - \langle f\rangle_{W,l}^{T''}(v)\big| < 4\rho.$$

This implies that the limit $\langle f\rangle_{W,l}(v)$ exists for every $v \in B(M, h^{s_0})$. Using (i) we obtain that $\langle f\rangle_{W,l}(\cdot) \in Lip_{\mathcal{C}}(h^{s_0})$.

(iii) This statement follows directly from (ii) since the family of functions $\{\langle f\rangle_{W,l}^T(v)\}$ is uniformly continuous on balls $B(R, h^{s_0})$ by (i) and each ball $B(M, h^s)$, $s > s_0$, is compact in h^{s_0}.

(iv) From the proof of (ii) we see that for any $\rho > 0$ and $v \in h^{s_0}$, there exists $T = T(W, \rho, v, \mathcal{C})$ such that if $T' \geqslant T$, then $|\langle f\rangle_{W,l}^{T'}(v) - \langle f\rangle_{W,l}(v)| \leqslant \rho$. This implies the assertion. □

We now give some examples of the limits $\langle f\rangle_{W,l}$.

Example 2.2. *If the vector W is non-resonant, i.e., non-trivial finite linear combinations of w_j's with integer coefficients do not vanish (this property holds for typical potentials $V(x)$, see [14]), then the set $S(K)$ reduces to one trivial resonance $e_l = (0, \dots, 0, 1, 0, \dots 0)$, where 1 stands on the l-th place (if $m < l$, then $S(K) = \emptyset$). Let $f(v)$ be any finite polynomial of v. We write it in the form*

$\sum_{k,l\in\mathbb{N}^\infty,|k|,|l|<\infty} f_{k,l}(I)v^k\bar{v}^l$, where $f_{k,l}$ are polynomials of I and finite vectors k,l are such that if $k_j\neq 0$, then $l_j=0$, and vice versa. Then $\langle f\rangle_{W,l}=f_{e_l,I}(I)v_l$.

Example 2.3. *If f is a linear functional, $f=\sum_{i=1}^\infty b_iv_i$, then for any $l\in\mathbb{N}$,*

$$\langle f\rangle_{W,l}=\sum_{i\in\mathcal{A}_l^1} b_iv_i,\quad \mathcal{A}_l^1=\{i\in\mathbb{N}: w_i-w_l=0\}.$$

If f is polynomial of v, e.g. $f=\sum_{i+j+m=k} a_{i,j,k}v_iv_jv_k$, then

$$\langle f\rangle_{W,l}=\sum_{(i,j,m)\in\mathcal{A}_l^3} a_{i,j,k}v_iv_jv_k,\quad \mathcal{A}_l^3=\{(i,j,m)\in\mathbb{N}^3: w_l-w_i-w_j-w_m=0\}.$$

We may also consider the averaging

$$\langle\langle f\rangle\rangle_W^T(v)=\frac{1}{T}\int_0^T f(\Phi_{-Wt}v)\,dt\,,\quad \langle\langle f\rangle\rangle_W(v)=\lim_{T\to\infty}\langle\langle f\rangle\rangle_W^T(v)\,. \tag{2.5}$$

Lemma 2.4. *Let $f\in Lip_{\mathcal{C}}(h^{s_0})$. Then*
a) for the averaging $\langle\langle\cdot\rangle\rangle_W$ hold natural analogies of all assertions of Lemma 2.1.
b) The function $\langle\langle f\rangle\rangle_W$ commutes with the transformations Φ_{Wt}, $t\in\mathbb{R}$.

Proof. To prove a) we repeat for the averaging $\langle\langle\cdot\rangle\rangle_W$ the proof of Lemma 2.1, replacing there w_l by 0. Assertion b) immediately follows from the formula for $\langle\langle f\rangle\rangle_W^T$ in (2.5). □

3. The effective equation

Let $V(x)\in C^n(T^d)$. As in the introduction, A_V is the operator $-\Delta+V$ and $\{\lambda_k,k\in\mathbb{N}\}$ are its eigenvalues.

The following result is well known, see Section 5.5.3 in [20].

Lemma 3.1. *If $f(x):\mathbb{C}\to\mathbb{C}$ is C^∞, then the mapping*

$$M_f: H^s\to H^s,\quad u\mapsto f(u),$$

is C^∞-smooth for $s>d/2$. Moreover, $M_f\in Lip_{\mathcal{C}_s}(H^s,H^s)$ for a suitable function $\mathcal{C}_s$.

Consider the map $P(v)$ defined in (1.5). From Lemma 3.1, we have

$$P(\cdot)\in Lip_{\mathcal{C}_s}(h^s,h^{s-1}),\quad \forall\, s\in(d/2+1,n]\,, \tag{3.1}$$

for some $\mathcal{C}_s$. We recall that Λ is the frequency vector of Eq. (1.2). For any $T\in\mathbb{R}$, we denote

$$\langle P\rangle_\Lambda^T(v):=(\langle P_k\rangle_{\Lambda,k}^T(v),k\in\mathbb{N})=\frac{1}{T}\int_0^T\Phi_{\Lambda t}P(\Phi_{-\Lambda t}v)dt\,,$$

and

$$R(v)=\langle P\rangle_\Lambda(v):=(\langle P_k\rangle_{\Lambda,k}(v),k\in\mathbb{N})\,.$$

Example 3.2. *If P is a diagonal operator, $P_k(v)=\gamma_kv_k$ for each k, where γ_k's are complex numbers, then in view of Example 2.3, $\langle P\rangle_\Lambda=P$.*

We have the following lemma:

Lemma 3.3. *(i) For every $d/2 < s_1 < s-1 \leqslant n-1$ and $M>0$, we have*

$$\Big|\langle P\rangle_\Lambda^T(v) - R(v)\Big|_{s_1} \to 0, \quad as \quad T\to\infty, \tag{3.2}$$

uniformly for $v \in B(M,h^s)$;
(ii) $R(\cdot) \in Lip_{\mathcal{C}_s}(h^s, h^{s-1})$, $s \in (d/2+1, n]$;
(iii) R commutes with $\Phi_{\Lambda t}$, for each $t\in\mathbb{R}$.

Proof. (i) There exists $M_1 > 0$, independent from v and T, such that

$$\big|\langle P\rangle_\Lambda^T(v) - R(v)\big|_{s-1} \leqslant M_1, \quad v \in B(M,h^s).$$

So for any $\rho > 0$ we can find $m_\rho > 0$ such that

$$\Big|(\mathrm{Id} - \Pi^{m_\rho})\big[\langle P\rangle_\Lambda^T(v) - R(v)\big]\Big|_{s_1} < \rho/2, \quad v \in B(M,h^s).$$

By Lemma 2.1(iii), there exists T_ρ such that for $T > T_\rho$,

$$\Big|\Pi^{m_\rho}\big[\langle P\rangle_\Lambda^T(v) - R(v)\big]\Big|_{s_1} < \rho/2, \quad v \in B(M,h^s).$$

Therefore if $T > T_\rho$, then

$$\big|\langle P\rangle_\Lambda^T(v) - R(v)\big|_{s_1} < \rho, \quad v \in B(M,h^s).$$

This implies the first assertion.

(ii) Using the fact that the linear maps $\Phi_{\Lambda t}$, $t\in\mathbb{R}$ are isometries in h^s, we obtain that for $T\in\mathbb{R}$ and $v', v'' \in B(M,h^s)$,

$$\big|\langle P\rangle_\Lambda^T(v') - \langle P\rangle_\Lambda^T(v'')\big|_{s-1} \leqslant \mathcal{C}_s(M)\,|v'-v''|_s\ .$$

Therefore

$$|R(v') - R(v'')|_{s-1} \leqslant \mathcal{C}_s(M)\,|v'-v''|_s\,, \quad v', v'' \in B(M,h^s).$$

This estimate, the convergence (3.2) and the Fatou lemma imply that R is a locally Lipschitz mapping with a required estimate for the Lipschitz constant. A bound on its norm may be obtained in a similar way, so the second assertion follows.

(iii) We easily verify that

$$\big|\langle P\rangle_\Lambda^{T+t}(v) - \Phi_{\Lambda t}\langle P\rangle_\Lambda^T(\Phi_{-\Lambda t}v)\big|_{s-1} \le 2\,\mathcal{C}_s(|v|_s)\frac{|t|}{|T+t|}\,.$$

Passing to the limit as $T\to\infty$ we recover (iii). ☐

Corollary 3.4. *For $d/2 < s_1 < s-1 \le n-1$ and any $v\in h^s$,*

$$\langle P\rangle_\Lambda^T(v) = R(v) + \varkappa(T;v),$$

where $|\varkappa(T;v)|_{s_1} \le \overline{\varkappa}(T;|v|_s)$. Here for each T, $\overline{\varkappa}(T;r)$ is an increasing function of r, and for each $r\ge 0$, $\overline{\varkappa}(T;r)\to 0$ as $T\to\infty$.

Example 3.5. *In the completely resonant case, when*

$$L_1 = \cdots = L_d = 2\pi \quad and \quad V = 0\,, \tag{3.3}$$

the frequency vector is $\Lambda = (|\mathbf{k}|^2, \mathbf{k}\in\mathbb{Z}^d)$. If $\mathcal{P}(u) = i|u|^2u$, then

$$P(v) = (P_{\mathbf{k}}(v), \mathbf{k}\in\mathbb{Z}^d), \quad v = (v_{\mathbf{k}}, \mathbf{k}\in\mathbb{Z}^d), \quad u = \sum_{\mathbf{k}\in\mathbb{Z}^d} v_{\mathbf{k}} e^{i\mathbf{k}\cdot x},$$

with

$$P_{\mathbf{k}}(v) = \sum_{\mathbf{k}_1-\mathbf{k}_2+\mathbf{k}_3=\mathbf{k}} iv_{\mathbf{k}_1}\bar{v}_{\mathbf{k}_2}v_{\mathbf{k}_3}, \quad \mathbf{k}\in\mathbb{Z}^d.$$

Therefore $\langle P\rangle_\Lambda = (\langle P_{\mathbf{k}}\rangle_{\Lambda,\mathbf{k}}, \mathbf{k}\in\mathbb{Z}^d)$, *with*

$$\langle P_{\mathbf{k}}\rangle_{\Lambda,\mathbf{k}} = \sum_{(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)\in Res(\mathbf{k})} iv_{\mathbf{k}_1}\bar{v}_{\mathbf{k}_2}v_{\mathbf{k}},$$

where $Res(\mathbf{k}) = \{(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) : |\mathbf{k}_1|^2 - |\mathbf{k}_2|^2 + |\mathbf{k}_3|^2 - |\mathbf{k}|^2 = 0\}$.

Lemma 3.3 implies that the effective equation (1.11) is a quasi-linear heat equation. So it is locally well-posed in the spaces h^s, $s\in(d/2+1,n]$.

4. Proof of the averaging theorem

In this section we will prove Theorem 1.1. We recall that $d/2+1 < s_* \le n$ and $s_1 < s_*$, where s_* is the number from Assumption A and n is a sufficiently big integer (the smoothness of the potential $V(x)$). Without loss of generality we assume that

$$s_1 > d/2+1 \quad \text{and} \quad s_1 > s_* - 2\,,$$

and that Assumption A holds with $T=1$.

Let $u^\epsilon(\tau,x)$ be the solution of Eq. (1.2) from Theorem 1.1,

$$\|u^\epsilon(0,x)\|_{s_*} \leqslant M_0\,,$$

and $v^\epsilon(\tau) = \Psi(u^\epsilon(\tau,\cdot))$. Then there exists $M_1 \geqslant M_0$ such that

$$v^\epsilon(\tau)\in B(M_1,h^{s_*}), \quad \tau\in[0,1]\,,$$

for each $\epsilon>0$. The constants in estimates below in this section may depend on M_1, and this dependence may be non-indicated.

Let

$$a^\epsilon(\tau) = \Phi_{\tau\epsilon^{-1}\Lambda}(v^\epsilon(\tau))$$

be the interaction representation of $v^\epsilon(\tau)$ (see Introduction),

$$a^\epsilon(0) = v(0) =: v_0\,.$$

For every $v=(v_k, k\in\mathbb{N})$, denote

$$\widehat{A}_V(v) = (\lambda_k v_k, k\in\mathbb{N}) = \Psi(A_V u)\,, \quad u = \Psi^{-1}v\,.$$

Then

$$\dot{a}^\epsilon(\tau) = -\mu\widehat{A}_V(a^\epsilon(\tau)) + Y\big(a^\epsilon(\tau),\epsilon^{-1}\tau\big)\,, \tag{4.1}$$

where

$$Y\big(a,t\big) = \Phi_{t\Lambda}\Big(P\big(\Phi_{-t\Lambda}(a)\big)\Big)\,. \tag{4.2}$$

Let $r\in(d/2+1,n]$. Since the operators $\Phi_{t\Lambda}$, $t\in\mathbb{R}$, define isometries of h^r, then, in view of (3.1), for any $t\in\mathbb{R}$ we have

$$Y(\cdot,t)\in Lip_{\mathcal{C}_r}(H^r,H^{r-1})\,. \tag{4.3}$$

For any $s\ge 0$ we denote by X^s the space

$$X^s = C([0,T],h^s)\,,$$

given the supremum-norm. Then

$$|a^\epsilon|_{X^{s_*}} \le M_1, \qquad |\dot{a}^\epsilon|_{X^{s_*-2}} \le C(M_1)\,. \tag{4.4}$$

Since for $0 \le \gamma \le 1$ we have

$$|v|_{\gamma(s_*-2)+(1-\gamma)s_*} \le |v|_{s_*-2}^{\gamma} |v|_{s_*}^{1-\gamma}$$

by the interpolation inequality, then in view of (4.4) for any $s_* - 2 < \bar{s} < s_*$ and $0 \le \tau_1 \le \tau_2 \le 1$ we have

$$|a^\epsilon(\tau_2) - a^\epsilon(\tau_1)|_{\bar{s}} \le C(M_1)^\gamma (\tau_2 - \tau_1)^\gamma (2M_1)^{1-\gamma}\,, \tag{4.5}$$

for a suitable $\gamma = \gamma(\bar{s}, s_*) > 0$, uniformly in ϵ.

Denote

$$\mathcal{Y}(v,t) = Y(v,t) - R(v).$$

Then by Lemma 3.3 relation (4.3) also holds for the map $v \mapsto \mathcal{Y}(v,t)$, for any t.

The following lemma is the main step of the proof.

Lemma 4.1. *For every $s' > d/2+1$, $s_* - 2 < s' < s_*$ we have*

$$\Big| \int_0^{\tilde{\tau}} \mathcal{Y}(a^\epsilon(\tau), \epsilon^{-1}\tau) d\tau \Big|_{s'} \leqslant \delta(\epsilon, M_1), \quad \forall\, \tilde{\tau} \in [0,1], \tag{4.6}$$

where $\delta(\epsilon, M_1) \to 0$ as $\epsilon \to 0$.

Proof. Below in this proof we write $a^\epsilon(\tau)$ as $a(\tau)$. We divide the time interval $[0,1]$ into subintervals $[b_{l-1}, b_l]$, $l = 1, \cdots, N$ of length $L = \epsilon^{1/2}$:

$$b_k = Lk \quad \text{for} \quad k = 0, \dots, N-1, \;\; b_N = 1\,, b_N - b_{N-1} \leqslant L\;,$$

where $N \leqslant 1/L + 1 \le 2/L$.

In virtue of (4.3) and Lemma 3.3 (ii),

$$\Big| \int_{b_{N-1}}^{b_N} \mathcal{Y}(a(\tau), \epsilon^{-1}\tau)\, d\tau \Big|_{s'} \le LC(s', s, M_1)\;. \tag{4.7}$$

Similar, if $\bar{\tau} \in [b_r, b_{r+1})$ for some $0 \le r < N$, then $|\int_{b_r}^{\bar{\tau}} \mathcal{Y}\, d\tau|_{s'}$ is bounded by the r.h.s. of (4.7).

Now we estimate the integral of $\mathcal{Y}$ over any segment $[b_l, b_{l+1}]$, where $l \le N-2$. To do this we write it as

$$\begin{aligned}\int_{b_l}^{b_{l+1}} \mathcal{Y}(a(\tau), \epsilon^{-1}\tau)\, d\tau = &\int_{b_l}^{b_{l+1}} \big(Y(a(b_l), \epsilon^{-1}\tau) - R(a(b_l))\big)\, d\tau \\ &+ \int_{b_l}^{b_{l+1}} \big(Y(a(\tau), \epsilon^{-1}\tau) - Y(a(b_l), \epsilon^{-1}\tau)\big)\, d\tau \\ &+ \int_{b_l}^{b_{l+1}} \big(R(a(b_l)) - R(a(\tau))\big)\, d\tau\,.\end{aligned}$$

In view of Lemma 3.3 and (4.5) the $h^{s'}$-norm of the second and third terms in the r.h.s. are bounded by $C(s', s, M_1)L^{1+\gamma}$. Since

$$\epsilon \int_0^{\epsilon^{-1}L} Y(a(b_l), \epsilon^{-1}b_l + s)\, ds = L\Phi_{\Lambda\epsilon^{-1}b_l} \frac{1}{L^{-1}} \int_0^{L^{-1}} \Phi_{\Lambda s} P\big(\Phi_{-\Lambda s}(\Phi_{-\Lambda\epsilon^{-1}b_l} a(b_l))\big) ds\,,$$

then using Corollary 3.4 and Lemma 3.3 (iii) we see that this equals $LR(a(b_l)) + \varkappa_1(L^{-1})$, where $|\varkappa_1(L^{-1})|_{s'} \le \overline{\varkappa}(L^{-1}; M_1)$ and $\overline{\varkappa} \to 0$ when $L^{-1} \to \infty$. We have arrived at the estimate

$$\Big| \int_{b_l}^{b_{l+1}} \mathcal{Y}(a(\tau), \epsilon^{-1}\tau)\, d\tau \Big|_{s'} \le L\Big(\overline{\varkappa}(L^{-1}; M_1) + CL^\gamma\Big)\,. \tag{4.8}$$

Since $N \le 2/L$ and $L = \epsilon^{1/2}$, then by (4.8) and (4.7) the l.h.s. of (4.6) is bounded by $2\overline{\varkappa}(\epsilon^{-1/2}; M_1) + C\epsilon^{\gamma/2} + C\epsilon^{1/2}$. It implies the assertion of the lemma. □

Consider the effective equation (1.11). By Lemma 3.3 this is the linear parabolic equation $\dot u - \Delta u + V(x)u = 0$, written in the v-variables, perturbed by a locally Lipschitz operator of order one. So its solution $\tilde a(\tau)$ such that $\tilde a(0) = v_0$ exists (at least) locally in time. Denote by $\tilde T$ the stopping time

$$\tilde T = \min\{\tau \in [0,1] : \ |\tilde a(\tau)|_{s_*} \geqslant M_1 + 1\},$$

where, by definition, $\min \emptyset = 1$.

Now consider the family of curves $a^\epsilon(\cdot) \in X^{s_*}$. In view of (4.4), (4.5) and the Arzelà-Ascoli theorem (e.g. see in [11]) this family is pre-compact in each space X^{s_1}, $s_1 < s_*$. Hence, for any sequence $\epsilon'_j \to 0$ there exists a subsequence $\epsilon_j \to 0$ such that

$$a^{\epsilon_j}(\cdot) \underset{\epsilon_j \to 0}{\longrightarrow} a^0(\cdot) \quad \text{in} \quad X^{s_1}.$$

By this convergence, (4.4) and the Fatou lemma,

$$|a^0(\tau)|_{s^*} \le M_1 \qquad \forall\, 0 \le \tau \le 1. \tag{4.9}$$

In view of Lemma 4.1, the curve $a^0(\tau)$ is a mild solution of Eq. (1.11) in the space h^{s_1}, i.e,

$$a(\tau) - a(0) = \int_0^\tau \big(-\mu \widehat{A}_V a(s) + R(a(s)\big)\, ds, \quad \forall\, 0 \le \tau \le 1$$

(the equality holds in the space $h^{s_1 - 2}$). So $a^0(\tau) = \tilde a(\tau)$ for $0 \le \tau \le \tilde T$. In view of (4.9) and the definition of the stopping time $\tilde T$ we see that $\tilde T = 1$. That is, $\tilde a \in X^{s_*}$ and

$$a^\epsilon(\cdot) \longrightarrow \tilde a(\cdot) \quad \text{in} \quad X^{s_1}, \tag{4.10}$$

where $\epsilon = \epsilon_j \to 0$. Since the limit $\tilde a$ does not depend on the sequence $\epsilon_j \to 0$, then the convergence holds as $\epsilon \to 0$.

Now we show that the convergence (4.10) holds uniformly for $v_0 \in B(M_0, h^{s_*})$. Assume the opposite. Then there exists $\delta > 0$, sequences $\tau_j \in [0,1], a_0^j \in B(M_0, h^{s_*})$, and $\epsilon_j \to 0$ such that if $a^{\epsilon_j}(\cdot)$ is a solution of (4.1) with initial data a_0^j and $\epsilon = \epsilon_j$, and $\tilde a^j(\cdot)$ is a solution of the effective equation (1.11) with the same initial data, then

$$|a^{\epsilon_j}(\tau_j) - \tilde a^j(\tau_j)|_{s_1} \geqslant \delta. \tag{4.11}$$

Using again the Arzelà-Ascoli theorem and (4.5), replacing the subsequence $\epsilon_j \to 0$ by a suitable subsequence, we have that

$$\begin{aligned}
&\tau_j \to \tau_0 \in [0,1],\\
&a_0^j \to a_0 \quad \text{in} \quad h^{s_1}, \quad \text{where} \quad a_0 \in h^{s_*},\\
&a^{\epsilon_j}(\cdot) \to a^0(\cdot) \quad \text{in} \quad X^{s_1},\\
&\tilde a^j(\cdot) \to \tilde a^0(\cdot) \quad \text{in} \quad X^{s_1}.
\end{aligned}$$

Clearly, $\tilde a^0(\cdot)$ is a solution of Eq. (1.11) with the initial datum a_0. Due to Lemma 4.1, $a^0(\cdot)$ is a mild solution of Eq. (1.11) with $a^0(0) = a_0$. Hence we have $a^0(\tau) = \tilde a^0(\tau)$, $\tau \in [0,1]$, particularly, $a^0(\tau_0) = \tilde a^0(\tau_0)$. This contradicts with (4.11), so the convergence (4.10) is uniform in $v_0 \in B(M_0, h^{s_*})$.

Since
$$|I(a)-I(\tilde a)|^{\sim}_{s_1}\le 4|a-\tilde a|_{s_1}(|a|_{s_1}+|\tilde a|_{s_1}),$$
then the convergence (4.10) implies the statement of Theorem 1.1.

5. The randomly forced case

We study here the effect of the addition a random forcing to Eq. (1.1). Namely, we consider equation (1.13). We suppose that
$$B_s=2\sum_{j=1}^{\infty}\lambda_j^{2s}b_j^2<\infty \qquad \text{for } s=s_*\in(d/2+1,n],$$
and impose a restriction on the nonlinearity $\mathcal{P}$ by assuming that there exists $\bar N\in\mathbb{N}$ and for each $s\in(d/2+1,n]$ there exists C_s such that
$$\|\mathcal{P}(\nabla u,u)\|_{s-1}\le C_s(1+\|u\|_s)^{\bar N}\,,\quad \forall\, u\in H^s \tag{5.1}$$
(this assumption holds e.g. if $\mathcal{P}(\nabla u,u)$ is a polynomial in $(u,\nabla u)$).

Passing to the slow time $\tau=\epsilon t$, Eq. (1.13) becomes (cf. (1.2))
$$\dot u+\epsilon^{-1}i(-\Delta+V(x))u=\mu\Delta u+\mathcal{P}(\nabla u,u)+\frac{d}{d\tau}\sum_{k=1}^{\infty}b_k\boldsymbol{\beta}_k e_k(x),\quad u=u(\tau,x), \tag{5.2}$$
which, in the v-variables, takes the form (cf. (1.4))
$$dv_k+\epsilon^{-1}i\lambda_k v_k\,d\tau=(-\mu\lambda_k v_k+P_k(v))\,d\tau+\sum_{l=1}^{\infty}\Psi_{kl}b_l d\boldsymbol{\beta}_l\ ,\quad k\in\mathbb{N}\ , \tag{5.3}$$
where we have denoted by $\{\Psi_{kl},k,l\ge 1\}$ the matrix of the operator Ψ (see (1.3)) with respect to the basis $\{e_k\}$ in H^0 and $\{\zeta_k\}$ in h^0. We assume

Assumption A$'$. *There exist $s_*\in(d/2+1,n]$ and an ϵ-independent $T>0$ such that for any $u_0\in H^{s_*}$, Eq. (5.2) has a unique strong solution $u(\tau,x)$, $0\le\tau\le T$, equal to u_0 at $\tau=0$. Furthermore, for each p there exists a $C=C_p(\|u_0\|_{s_*},B_{s_*},T)$ such that*
$$\mathbf{E}\sup_{0\le\tau\le T}\|u(\tau)\|_{s_*}^p\le C\,. \tag{5.4}$$

Remark 5.1. *The Assumption A$'$ is not too restrictive. In particular, in* [14] *it is verified for equations* (1.13) *if $\mu>0$ and $\mathcal{P}(u)=-u+zf_p(|u|^2)u$, where $f_p(r)$ is a smooth function, equal $|r|^p$ for $|r|\ge 1$, and $\operatorname{Im}z\le 0,\operatorname{Re}z\le 0$. The degree p is any real number if $d=1,2$ and $p<2/(d-2)$ if $d\ge 3$.*

Under this assumption, a result analogous to Theorem 1.1 holds. Namely, the limiting behaviour of the action variables I_k (see (1.6)) is described by the stochastically forced effective equation (cf. (1.11))
$$d\tilde a_k=(-\mu\lambda_k\tilde a_k+R_k(\tilde a))\,d\tau+\sum_{l=1}^{\infty}B_{kl}d\boldsymbol{\beta}_l\ ,\quad k\in\mathbb{N}\ , \tag{5.5}$$
where we have defined $\{B_{kr},k,r\ge 1\}$ as the principal square root of the real matrix
$$A_{kr}=\begin{cases}\sum_l b_l^2\Psi_{kl}\Psi_{rl} & \text{if } \lambda_k=\lambda_r\,,\\ 0 & \text{else}\,,\end{cases}\quad . \tag{5.6}$$
which defines a nonnegative selfadjoint compact operator in h^0. Note that since R is locally Lipschitz by Lemma 3.3, then strong solutions for (5.5) exist and are

unique till the stopping time $\tau_K = \inf\{\tau \geq 0 : |\tilde{a}(\tau)|_{s_*} = K\}$, where K is any positive number.

In the theorem below $v^\epsilon(\tau)$ denotes a solution of (5.3) with the initial value $v_0 \in h^{s_*}$.

Theorem 5.2. *If Assumption A′ holds, there exists a unique strong solution $\tilde{a}(\tau)$, $0 \leq \tau \leq T$, of equation (5.5) such that $\tilde{a}(0) = v_0 = \Psi(u_0) \in h^{s_*}$, and*

$$\mathcal{D}\left(I(v^\epsilon(\tau))\right) \rightharpoonup \mathcal{D}\left(I(\tilde{a}(\tau))\right) \quad \text{as } \epsilon \to 0 \ ,$$

in $C([0,T], h_I^{s_1})$, for any $s_1 < s_*$.

In the theorem's assertion and below the arrow $\rightharpoonup$ stands for the weak convergence of measures. Let us assume further:

Assumption B′. *i) Eq.* (1.13) *has a unique strong solution $u(\tau), u(0) = u_0 \in H^{s_*}$, defined for $\tau \geq 0$, and*

$$\mathbf{E} \sup_{\theta \leq \tau \leq \theta+1} \|u(\tau)\|_{s_*}^p \leq C \quad \text{for any } \theta \geq 0 \,, \tag{5.7}$$

where $C = C(\|u_0\|_{s_*}, B_{s_*})$.

ii) Eq. (5.5) *has a unique stationary measure μ^0 and is mixing.*

Remark 5.3. *The assumption i) is fulfilled, for example, for equations, discussed in Remark 5.1. Assumption ii) holds trivially if for a.e. realisation of the random force any two solutions of Eq.* (5.5) *converge exponentially fast.*[1] *For less trivial examples, corresponding to perturbations of linear systems with non-resonant or completely resonant spectra, see* [14, 15].

Assumption B′ i) and the Bogolyubov-Krylov argument, applies for solutions, starting from 0, imply that Eq. (1.13) has a stationary measure μ^ϵ, supported by the space H^{s_*}, and inheriting estimates (5.7).

Theorem 5.4. *Let us suppose that Assumptions A′ and B′ hold. Then*

$$\lim_{\epsilon \to 0} \mu^\epsilon = \mu^0 \ , \tag{5.8}$$

weakly in h^{s_1}, for any $s_1 < s_*$. The measure μ^0 is invariant with respect to transformations $\Phi_{t\Lambda}$, $t \in \mathbb{R}$. If, in addition, (1.13) *is mixing and μ^ϵ is its unique stationary measure, then for any solution $u^\epsilon(t)$ of* (1.13) *with ϵ-independent initial data $u_0 \in H^{s_*}$, we have*

$$\lim_{\epsilon \to 0} \lim_{t \to \infty} \mathcal{D}\left(v^\epsilon(t)\right) = \mu^0 \,,$$

where $v^\epsilon(t) = \Psi\left(u^\epsilon(t)\right)$.

For examples of mixing equations (1.13) see [14] and references in that work. In particular, (1.13) is mixing if $\mathcal{P}(u, \nabla u) = \mathcal{P}(u)$ is a smooth function such that all its derivatives are bounded uniformly in u, cf. Remark 5.3.

For the case when the spectrum Λ is non-resonant (see Example 2.2) or is completely resonant, i.e. (3.3) holds, the theorem was proved in [14, 15].

The proofs of Theorem 5.2 and 5.4 closely follow the arguments in [14, 15, 16]. Proof of Theorem 5.4, in addition, uses some technical ideas from [4] (see there Corollary 4.2). The proofs are given, respectively, in Section 6 and Section 7.

---

[1] This is fulfilled, for example, if i) holds and $\mathcal{P}(u) = -u + \mathcal{P}_0(u)$, where the Lipschitz constant of $\mathcal{P}_0$ is less than one.

6. Proof of Theorem 5.2

As in the proof of Theorem 1.1, let us assume, without loss of generality, that $T = 1$, $s_1 > d/2 + 1$ and $s_1 > s_* - 2$ (recall that $s_1 < s_*$ and $s_* \in (d/2 + 1, n]$).

Following the suite of [15] (see also [16]) we pass once again to the a-variables, defined in (1.8)). In view of (5.3), they satisfy the system (cf. (1.10))

$$da_k = \big(-\mu\lambda_k a_k + Y_k(a, \epsilon^{-1}\tau)\big)\, d\tau + e^{i\epsilon^{-1}\lambda_k\tau} \sum_l \Psi_{kl} b_l d\boldsymbol{\beta}_l, \quad k \in \mathbb{N}\,, \tag{6.1}$$

where Y is defined in (4.2). For any p we denote

$$X^p = C([0,1], h^p)\,, \qquad X_I^p = C([0,1], h_I^p)\,.$$

Let a^ϵ be a solution of (6.1) such that $a^\epsilon(0) = v_0 = \Psi(u_0) \in h^{s_*}$; we will often write a for a^ϵ to shorten notation. Denote the white noise in (6.1) as $\dot\zeta(t,x)$ and denote $U_1(\tau) = Y(a(\tau), \epsilon^{-1}\tau)$, $U_2(\tau) = -\widehat{A}_V a(\tau)$. Then

$$\dot a - \dot\zeta = U_1 + U_2\,.$$

In view of (5.1), $\|U_1\|_{s_*-1} = |P(v)|_{s_*-1} \le C(1 + \|u(\tau)\|_{s_*}^{\bar N})$. So, by (5.4),

$$\mathbf{E}\int_\tau^{(\tau+\tau')\wedge 1} \|U_1\|_{s_*-1}\, dt \le C \int_\tau^{(\tau+\tau')\wedge 1} \mathbf{E}C(1 + \|u(t)\|_{s_*}^{\bar N})\, dt \le C(\|u_0\|_{s_*}, B_{s_*})\tau'\,,$$

for any $\tau \in [0,1]$ and $\tau' > 0$. Similar,

$$\mathbf{E}\int_\tau^{(\tau+\tau')\wedge 1} \|U_2\|_{s_*-2}\, dt \le \mu C \mathbf{E}\int_\tau^{(\tau+\tau')\wedge 1} \|u\|_{s_*} \le \mu C(\|u_0\|_{s_*}, B_{s_*})\tau'\,.$$

Hence, there exists $\gamma > 0$ such that

$$\mathbf{E}\,\|(a-\zeta)((\tau+\tau')\wedge 1) - (a-\zeta)(\tau)\|_{s_1} \le C(\|u_0\|_{s_*}, B_{s_*}){\tau'}^{\gamma}\,,$$

in virtue of the interpolation and Hölder inequalities (cf. (4.5)). It is classical that

$$\mathbf{P}\{\|\zeta\|_{C^{1/3}([0,1],h^{s_1})} \le R_3\} \to 1 \quad \text{as} \quad R_3 \to \infty\,.$$

In view of what was said, for any $\delta > 0$ there is a set $Q_\delta^1 \subset X^{s_1}$, formed by equicontinuous functions, such that

$$\mathbf{P}\{a^\epsilon \in Q_\delta^1\} \ge 1 - \delta\,,$$

for each ϵ. By (5.4),

$$\mathbf{P}\{\|a^\epsilon\|_{X^{s_*}} \ge C\delta^{-1}\} \le \delta\,,$$

for a suitable C, uniformly in ϵ. Consider the set

$$Q_\delta = \left\{a^\epsilon \in Q_\delta^1 : \|a\|_{X^{s_*}} \le C\delta^{-1}\right\}\,.$$

Then $\mathbf{P}\{a^\epsilon \in Q_\delta\} \ge 1 - 2\delta$, for each ϵ. By this relation and the Arzelà-Ascoli theorem (e.g., see [11], §8), the set of laws $\{\mathcal{D}(a^\epsilon(\cdot)),\ 0 < \epsilon \le 1\}$, is tight in X^{s_1}. So by the Prokhorov theorem there is a sequence $\epsilon_l \to 0$ and a Borel measure $\mathcal{Q}^0$ on X^{s_1} such that

$$\mathcal{D}(a^{\epsilon_l}(\cdot)) \rightharpoonup \mathcal{Q}^0 \quad \text{as} \quad \epsilon_l \to 0\,. \tag{6.2}$$

Accordingly, due to (1.9), for actions of solutions v^ϵ we have the convergence

$$\mathcal{D}\left(I\left(v^{\epsilon_l}(\cdot)\right)\right) \rightharpoonup I \circ \mathcal{Q}^0 \quad \text{as} \quad \epsilon_l \to 0\,, \tag{6.3}$$

in $X_I^{s_1}$.

Theorem 5.2 follows then as a simple corollary from

Proposition 6.1. *There exists a unique weak solution $a(\tau)$ of the effective equation* (5.5) *such that $\mathcal{D}(a) = \mathcal{Q}^0$, $a(0) = v^0$ a.s.; and the convergences* (6.2) *and* (6.3) *hold as $\epsilon \to 0$.*

Proof. The proof follows the Khasminski scheme (see [12, 6]), as expounded in [15]. Namely, we show that the limiting measure $\mathcal{Q}^0$ is a martingale solution of the limiting equation, which turns out to be exactly the equation (5.5). Since the equation has a unique solution, then the convergences (6.2), (6.3) hold as $\epsilon \to 0$.

For $\tau \in [0,1]$ consider the processes

$$N_k^{\epsilon_l} = a_k^{\epsilon_l}(\tau) - \int_0^\tau \left(-\mu\lambda_k a_k^{\epsilon_l}(s) + R_k(a^{\epsilon_l}(s))\right) ds \ , \quad k \geq 1$$

(cf. Eq. (5.5)). Due to (6.1) we write $N_k^{\epsilon_l}$ as

$$N_k^{\epsilon_l}(\tau) = \widetilde{N}_k^{\epsilon_l}(\tau) + \overline{N}_k^{\epsilon_l}(\tau) \ ,$$

where $\widetilde{N}_k^{\epsilon_l}(\tau) = a^{\epsilon_l}(\tau) - \int_0^\tau (-\mu\lambda_k a^{\epsilon_l}(s) + Y_k(a^{\epsilon_l}(s), \epsilon_l^{-1}s))ds$ is a $\mathcal{Q}^0$ martingale and the disparity $\overline{N}_k^{\epsilon_l}$ is

$$\overline{N}_k^{\epsilon_l}(\tau) = \int_0^\tau \mathcal{Y}_k(a^{\epsilon_l}(s), \epsilon_l^{-1}s)ds$$

(as before, $\mathcal{Y}(a,t) = Y(a,t) - R(a)$).

The key point is then a stochastic counterpart of Lemma 4.1, which is proved below:

Lemma 6.2. *For every $k \in \mathbb{N}$, $\mathbf{E}\,\mathfrak{A}_k^\epsilon \to 0$ as $\epsilon \to 0$, where*

$$\mathfrak{A}_k^\epsilon = \max_{0 \leq \tilde\tau \leq 1} \left| \int_0^{\tilde\tau} \mathcal{Y}_k(a^\epsilon(\tau), \epsilon^{-1}\tau)d\tau \right| .$$

This lemma and the convergence (6.2) imply that the processes

$$N_k(\tau) = a_k(\tau) - \int_0^\tau \left(-\mu\lambda_k a_k + R_k(a)\right) ds \ , \quad k \geq 1 \ ,$$

are $\mathcal{Q}^0$ martingales, considered on the probability space $(\Omega = X^{s_1}, \mathcal{F}, Q^0)$ ($\mathcal{F}$ is the Borel sigma-algebra), given the natural filtration $(\mathcal{F}_\tau, 0 \leq \tau \leq 1)$. For details see [17], Proposition 6.3).

Consider then the diffusion matrix $\{\mathcal{A}_{kr}, k, r \geq 1\}$ for the system (6.1), i.e.,

$$\mathcal{A}_{kr} = \exp(i\epsilon^{-1}\tau(\lambda_k - \lambda_r)) \sum_{l=1}^\infty b_l^2 \Psi_{kl} \bar\Psi_{rl} \ .$$

Clearly, $\int_0^{\tilde\tau} \mathcal{A}_{kr} d\tau \to A_{kr}\tilde\tau$, as $\epsilon \to 0$, where A denotes the diffusion matrix for the system (5.5) (cf. (5.6)). Similar to Lemma 6.2, we also find that

$$\mathbf{E} \max_{0 \leq \tilde\tau \leq 1} \left| \int_0^{\tilde\tau} \mathcal{Y}_k(a^\epsilon(\tau), \epsilon^{-1}\tau)d\tau \right|^2 \to 0 \quad \text{as } \epsilon \to 0 \ .$$

Then, using the same argument as before, we see that the processes

$$N_k(\tau)N_r(\tau) - A_{kr}\tau = \left(\widetilde{N}_k \widetilde{N}_r - \int_0^\tau \mathcal{A}_{kr} ds\right) + \left(\overline{N}_k \overline{N}_r + \overline{N}_k \widetilde{N}_r + \widetilde{N}_k \overline{N}_l - \int_0^\tau (\mathcal{A}_{kr} - A_{kr})\, ds\right)$$

are $\mathcal{Q}^0$ martingales. That is, $\mathcal{Q}^0$ is a solution of the martingale problem with the drift R and the diffusion A (see [21]), so the assertion follows. $\square$

Proof of Lemma 6.2. We adopt a convenient notation from our previous publications. Namely, we denote by $\varkappa(r)$ various functions of r such that $\varkappa \to 0$ as $r \to \infty$. We write $\varkappa(r; M)$ to indicate that $\varkappa(r)$ depends on a parameter M. Besides for events Q and O and a random variable f we write $\mathbf{P}_O(Q) = \mathbf{P}(O \cap Q)$ and $\mathbf{E}_O(f) = \mathbf{E}(\chi_O f)$.

The constants below may depend on k, but this dependence is not indicated since k is fixed through the proof. By $M \geq 1$ we denote a constant which will be specified later. Denote by $\Omega_M = \Omega_M^\epsilon$ the event

$$\Omega_M = \left\{ \sup_{0 \leq \tau \leq 1} |a^\epsilon(\tau)|_{s_*} \leq M \right\} .$$

Then, by (5.4),

$$\mathbf{P}(\Omega_M^c) \leq \varkappa(M). \tag{6.4}$$

In view of Lemma 3.3 (ii) and (5.1), for any $t \in [0, \epsilon^{-1}]$ and any $a \in h^{s_*}$ the difference $\mathcal{Y} = Y - R$ satisfies

$$|\mathcal{Y}_k(a,t)| \leq |Y_k(a,t)| + |R_k(a)| \leq |P_k(v)| + |R_k(a)| \leq C(1 + |a|_{s_*})^{\bar{N}} . \tag{6.5}$$

Using this and (6.4) we get

$$\begin{aligned} \mathbf{E}_{\Omega_M^c} \mathfrak{A}_k^\epsilon &\leq \int_0^1 \mathbf{E}_{\Omega_M^c} |\mathcal{Y}_k(a(\tau), \epsilon^{-1}\tau)| d\tau \\ &\leq C \left(\mathbf{P}(\Omega_M^c)\right)^{1/2} \int_0^1 \left(\mathbf{E}(1 + |a|_{s_*})^{2\bar{N}}\right)^{1/2} d\tau \leq \varkappa(M) . \end{aligned} \tag{6.6}$$

To estimate $\mathbf{E}_{\Omega_M} \mathfrak{A}_k^\epsilon$, as in Lemma 4.1 we consider a partition of $[0,1]$ by the points

$$b_n = nL, \quad 0 \leq n \leq N-1, \ \ b_{N-1} \geq 1 - L, \ b_N = 1 , \qquad L = \epsilon^{1/2} ,$$

$N \sim 1/L$. Let us denote

$$\eta_l = \int_{b_l}^{b_{l+1}} \mathcal{Y}_k(a(\tau), \epsilon^{-1}\tau) d\tau \ , \quad 0 \leq l \leq N-1 .$$

Since for $\omega \in \Omega_M$ and any $\tau' < \tau''$ such that $\tau'' - \tau' \leq L$, in view of (6.5) we have $\left| \int_{\tau'}^{\tau''} \mathcal{Y}_k(a(\tau), \epsilon^{-1}\tau) d\tau \right| \leq LC(M)$, then

$$\mathbf{E}_{\Omega_M} \mathfrak{A}_k^\epsilon \leq LC(M) + \mathbf{E}_{\Omega_M} \sum_{l=0}^{N-1} |\eta_l| . \tag{6.7}$$

Let us fix any $\bar{s} > d/2 + 1$, $s_* - 2 < \bar{s} < s_*$, sufficiently small $\gamma > 0$, and consider the event

$$\mathcal{F}_l = \left\{ \sup_{b_l \leq \tau \leq b_{l+1}} |a^\epsilon(\tau) - a^\epsilon(b_l)|_{\bar{s}} \geq L^\gamma \right\} .$$

By the equicontinuity of the processes $\{a^\epsilon(\tau)\}$ on suitable events with arbitrarily close to one ϵ-independent probability (as shown above), the probability of $\mathbf{P}(\mathcal{F}_l)$

goes to zero with L, uniformly in l and ϵ. Since $|\eta_l| \le C(M)L$ for $\omega \in \Omega_M$ and each l, then

$$\sum_{l=0}^{N-1} \big| \mathbf{E}_{\Omega_M} |\eta_l| - \mathbf{E}_{\Omega_M \setminus \mathcal{F}_l} |\eta_l| \big| \le C(M)L \sum_{l=0}^{N-1} \mathbf{P}_{\Omega_M}(\mathcal{F}_l) \le C(M)\varkappa(L^{-1}) \ , \tag{6.8}$$

and it remains to estimate $\sum_l \mathbf{E}_{\Omega_M \setminus \mathcal{F}_l} |\eta_l|$.

We have

$$\begin{aligned} |\eta_l| \le & \left| \int_{b_l}^{b_{l+1}} \big(\mathcal{Y}_k(a(\tau), \epsilon^{-1}\tau) - \mathcal{Y}_k(a(b_l), \epsilon^{-1}\tau) \big) \, d\tau \right| \\ & + \left| \int_{b_l}^{b_{l+1}} \big(\mathcal{Y}_k(a(b_l), \epsilon^{-1}\tau) \big) \, d\tau \right| =: \Upsilon_l^1 + \Upsilon_l^2 \ . \end{aligned}$$

By (3.1) and Lemma 3.3 (ii), in Ω_M the following inequality hold:

$$\big| \mathcal{Y}_k(a(\tau), \epsilon^{-1}\tau) - \mathcal{Y}_k(a(b_l), \epsilon^{-1}\tau) \big| \le C(M) \, |a(\tau) - a(b_l)|_{\bar{s}} \ .$$

So that, by the definition of $\mathcal{F}_l$,

$$\sum_l \mathbf{E}_{\Omega_M \setminus \mathcal{F}_l} \Upsilon_l^1 \le L^\gamma C(M) = \varkappa(\epsilon^{-1}; M) \ . \tag{6.9}$$

It remains to estimate the expectation of $\sum \Upsilon_l^2$. In view of (4.8) (with $M_1 = M$) we have

$$\sum_l \mathbf{E}_{\Omega_M \setminus \mathcal{F}_l} \Upsilon_l^2 \le NL\varkappa_1(\epsilon^{-1}; M) = \varkappa(\epsilon^{-1}; M). \tag{6.10}$$

Now the inequalities (6.6)–(6.10) jointly imply that

$$\mathbf{E} \, \mathfrak{A}_k^\epsilon \le \varkappa(M) + \varkappa(\epsilon^{-1}; M) \ .$$

Choosing first M large and then ϵ small we make the r.h.s. arbitrarily small. This proves the lemma. □

Lemma 6.2 estimates integrals of the differences

$$e^{i\epsilon^{-1}\tau\lambda_k} P_k \big(\Phi_{-\epsilon^{-1}\tau\lambda_k}(a^\epsilon(\tau)\big) - \langle P \rangle_{\Lambda,k}(a^\epsilon(\tau)) \, .$$

Similar result holds if we replace the averaging $\langle \cdot \rangle_{\Lambda,k}$ by $\langle\langle \cdot \rangle\rangle_\Lambda$ and the function P_k by any Lipschitz function:

Lemma 6.3. *Let $f \in Lip_1(h^{s_1}) =: Lip_1$ (i.e., f is a bounded Lipschitz function on h^{s_1}). Then*

i) $\mathbf{E} \int_0^1 \big(f(\Phi_{-\tau\epsilon^{-1}\Lambda} a^\epsilon(\tau)) - \langle\langle f \rangle\rangle_\Lambda (a^\epsilon(\tau)) \big) \, d\tau \to 0$ *as* $\epsilon \to 0$;

ii) if in i) f is replaced by $f^\theta = f \circ \Phi_\theta$, $\theta \in \mathbb{T}^\infty$, then the rate of convergence does not depend on θ.

Proof. To get i) we literally repeat the proof of Lemma 6.2, using Lemma 2.4 instead of Lemma 2.1. The assertion ii) follows from Lemma 2.4 and item (iv) of Lemma 2.1. □

7. Proof of Theorem 5.4

Let $v^\epsilon(\tau)$, $0 \le \tau \le 1$, be a stationary solution for Eq. (1.13) such that $\mathcal{D}(v^\epsilon(\tau)) \equiv \mu^\epsilon$, and let $a^\epsilon(\tau) = \Phi_{\epsilon^{-1}\Lambda\tau} v^\epsilon(\tau)$ be its interaction representation. Since v inherits the a-priori estimate (5.7) (with $u_0 = 0$), then an analogy of the convergence (6.2) holds for a suitable sequence $\epsilon_l \to 0$. The argument from the proof of Proposition 6.1 applies and imply that

$$\mathcal{D}(a^{\epsilon_l}(\cdot)) \rightharpoonup \mathcal{D}(a^0(\cdot)) \quad \text{in} \quad X^{s_1} \quad \text{as} \quad \epsilon_l \to 0\,, \tag{7.1}$$

where a^0 is a weak solution of (5.5). We may also assume that

$$\mu^{\epsilon_l} \rightharpoonup \bar\mu^0 \quad \text{in} \quad h^{s_1}\,, \tag{7.2}$$

for some measure $\bar\mu^0$.

Let us take any $f \in Lip_1(h^{s_1})$. Then

$$\mathbf{E}\int_0^1 f(v^\epsilon(\tau))\,d\tau = \mathbf{E}\int_0^1 f(\Phi_{-\epsilon^{-1}\Lambda\tau} a^\epsilon(\tau))\,d\tau\,.$$

Applying to the second integral Lemma 6.3 we find that

$$\int_0^1 \mathbf{E} f(v^\epsilon(\tau))\,d\tau = \int_0^1 \mathbf{E}\langle\langle f\rangle\rangle_\Lambda(a^\epsilon(\tau))\,d\tau + \varkappa(\epsilon^{-1})\,. \tag{7.3}$$

Since the function $\langle\langle f\rangle\rangle_\Lambda$ is invariant with respect to transformations $\Phi_{\Lambda t}$, $t \in \mathbb{R}$ (see item b) of Lemma 2.4), then $\langle\langle f\rangle\rangle_\Lambda(a^\epsilon(\tau)) = \langle\langle f\rangle\rangle_\Lambda(v^\epsilon(\tau))$. So both integrands in (7.3) are independent from τ, and

$$\mathbf{E} f(v^\epsilon(\tau)) = \mathbf{E}\langle\langle f\rangle\rangle_\Lambda(a^\epsilon(\tau)) + \varkappa(\epsilon^{-1}) \quad \forall\,\tau\,. \tag{7.4}$$

Now let us take for f the function $\tilde f = \tilde f_{\epsilon^{-1}\tau} = f \circ \Phi_{\epsilon^{-1}\Lambda\tau}$ (which also belongs to $Lip_1(h^{s_1})$). Then

$$\mathbf{E} f(a^\epsilon(\tau)) = \mathbf{E}\tilde f(v^\epsilon(\tau)) = \mathbf{E}\langle\langle \tilde f\rangle\rangle_\Lambda(a^\epsilon(\tau)) + \varkappa(\epsilon^{-1}) = \mathbf{E}\langle\langle f\rangle\rangle_\Lambda(a^\epsilon(\tau)) + \varkappa(\epsilon^{-1})\,,$$

where $\varkappa$ may be chosen the same for all functions $\tilde f$ in view of Lemma 6.3 i). Comparing this with (7.4) and using (7.2) we find that

$$\mathbf{E} f(a^{\epsilon_l}(\tau)) \rightharpoonup \langle f, \bar\mu^0\rangle \quad \text{as} \quad \epsilon_l \to 0,$$

for each τ. Therefore, in virtue of (7.1), $\mathcal{D}(a^0(\tau)) \equiv \bar\mu^0$. So $a^0(\tau)$ is a stationary solution for (5.5), and $\bar\mu^0$ is a stationary measure for this equation. Since the latter is unique, $\bar\mu^0 \equiv \mu^0$, and (7.2) implies the convergence (5.8).

Replacing in (7.4) f by $\tilde f_t$ and using Lemma 2.4 b) we see that

$$\langle f, \Phi_{\Lambda t} \circ \mu^\epsilon\rangle = \langle \tilde f_t, \mu^\epsilon\rangle = \langle f, \mu^\epsilon\rangle + \varkappa(\epsilon^{-1}).$$

Passing to the limit as $\epsilon \to 0$ we get the claimed invariance of the measure μ^0. Finally, the last assertion immediately follows from (5.8). $\square$

Appendix A

Consider the CGL equation (1.14), where $\mathcal{P} : \mathbb{C}^{d(d+1)/2+d+1} \times T^d \to \mathbb{C}$ is a C^∞-smooth function. We write it in the v-variables and slow time $\tau = \epsilon t$:

$$\dot{v}_k + \epsilon^{-1} i\lambda_k v_k = P_k(v), \quad k \in \mathbb{N},$$

where

$$P(v) := (P_k(v), k \in \mathbb{N}) = \Psi(\mathcal{P}(\nabla^2 u, \nabla u, u, x)), \quad u = \Psi^{-1} v,$$

and introduce the effective equation

$$\dot{\tilde{a}} = \langle P \rangle_\Lambda(\tilde{a}). \tag{A.1}$$

By Lemma 3.1 P defines smooth locally Lipschitz mappings $h^s \to h^{s-2}$ for $s > 2 + d/2$. So by a version of Lemma 3.3, $\langle P \rangle_\Lambda \in Lip_{loc}(h^s; h^{s-2})$ for $s > 2 + d/2$. Assume that

Assumption E: *There exists $s_0 \in (d/2, n]$ such that the effective equation* (A.1) *is locally well posed in the Hilbert spaces h^s, with $s \in [s_0, n] \cap \mathbb{N}$.*

Let $u^\epsilon(t, x)$ be a solution of Eq. (1.14) with initial datum $u_0 \in H^s$, $v^\epsilon(\tau) = \Psi(u(\epsilon^{-1}\tau, x))$, and $\tilde{a}(\tau)$ be a solution of Eq. (A.1) with initial datum $\Psi(u_0)$. Then we have the following result:

Theorem A.1. *If Assumptions A and E hold and $s > \max\{s_0 + 2, d/2 + 4\}$, then the solution of the effective equation exists for $0 \leqslant \tau \leqslant T$, and for any $s_1 < s$ we have*

$$I(v^\epsilon(\cdot)) \xrightarrow[\epsilon \to 0]{} I(\tilde{a}(\cdot)) \quad in \quad C([0, T], h_I^{s_1}).$$

The proof of this theorem follows that of Theorem 1.1, with slight modifications. Cf. [8], where the result is proven for the non-resonant case.

References

[1] V. Arnold, V. V. Kozlov, and A. I. Neistadt. *Mathematical Aspects of Classical and Celestial Mechanics*. Springer, Berlin, third edition, 2006.

[2] D. Bambusi. Galerkin averaging method and Poincaré normal form for some quasilnear PDEs. *Ann. Scuola Norm. Sup. Pisa C1. Sci.*, pages 669–702, 2005.

[3] N. N. Bogoljubov and J. A. Mitropol'skij. *Asymptotic Methods in the Theory of Non-Linear Oscillations*. Gordon and Breach, New York, 1961.

[4] A Dymov. Nonequilibrium statistical mechanics of weakly stochastically perturbed system of oscillators. *Preprint*, 2015. arXiv:1501.04238.

[5] E. Faou, P. Germain, and Z. Hani. The weakly nonlinear large box limit of the 2D cubic nonlinear Schrödinger equation. *Preprint*, 2013.

[6] M. I. Freidlin and A. D. Wentzell. Averaging principle for stochastic perturbations of multi-frequency systems. *Stochastics and Dynamics*, 3:393–408, 2003.

[7] P. Gérard and S. Grellier. Effective integrable dynamics for a certain nonlinear wave eqaution. *Analysis and PDE*, 5:1139–1154, 2012.

[8] G. Huang. An averaging theorem for nonlinear Schrödinger equations with small nonlinearities. *DCDS-A*, 34(9):3555–3574, 2014.

[9] G. Huang. Long-time dynamics of resonant weakly nonlinear CGL equations. *JDDE*, pages 1–13, 2014. doi:10.1007/s10884-014-9391-0.

[10] G. Huang and S.B. Kuksin. KdV equation under periodic boundary conditions and its perturbations. *Nonlinearity*, 27:1–28, 2014.

[11] J.L. Kelley and I. Namioka. *Linear topological spaces*. Springer-Verlag, New York-Heidelberg, 1976.

[12] R. Khasminski. On the averaging principle for Ito stochastic differential equations. *Kybernetika*, 4:260–279, 1968. (in Russian).

[13] S. Kuksin. Damped-driven KdV and effective equations for long-time behavior of its solutions. *GAFA*, 20:1431–1463, 2010.

[14] S. Kuksin. Weakly nonlinear stochastic CGL equations. *Annales de l'insitut Henri Poincaré-Probabilité et Statistiques*, 49(4):1033–1056, 2013.

[15] S. Kuksin and A. Maiocchi. Resonant averaging for weakly nonlinear stochastic Schrödinger equations. *Preprint*, 2013. arXiv:1309.5022.

[16] S. Kuksin and A. Maiocchi. The limit of small Rossby numbers for randomly forced quasi-geostrophic equation on β-plane. *Nonlinearity*, 28, 2015.

[17] S. Kuksin and A. Piatnitski. Khasminskii-Whitham averaging for randomly perturbed KdV equation. *J.Math. Pures Appl.*, 89:400–428, 2008.

[18] P. Lochak and C. Meunier. *Multiphase Averaging for Classical Systems*. Springer-Verlag, New York–Berlin–Heidelberg, 1988.

[19] S. Nazarenko. *Wave Turbulence*. Springer, Berlin, 2011.

[20] T. Runst and W. Sickel. *Sobolev spaces of fractional order, Nemytskij operators, and non-linear partial differential equations*, volume 3. de Gruyter, 1996.

[21] D. Stroock and S.R.S. Varadhan. *Multidimensional Diffusion Processes*. Springer-Verlag, New York–Berlin–Heidelberg, 1979.

[22] A. Zygmund. *Trigonometric Series*, volume 2. Cambridge University Press, Cambridge, 3 edition, 2002.