NNT : 2016SACLX029

THÈSE DE DOCTORAT DE L'UNIVERSITÉ PARIS-SACLAY

préparée à

École polytechnique

ÉCOLE DOCTORALE N°574

École doctorale de mathématiques Hadamard (EDMH)

Spécialité de doctorat : Mathématiques fondamentales

par

Jacek JENDREJ

Sur la dynamique d'équations des ondes avec non-linéarité énergie critique focalisante

Thèse présentée et soutenue à Palaiseau, le 11 juillet 2016

Après avis des rapporteurs :

M. Fabrice PLANCHON	Université de Nice
M. Wilhelm SCHLAG	University of Chicago

Composition du jury :

M. Carlos KENIG	Président	University of Chicago
M. Fabrice PLANCHON	Rapporteur	Université de Nice
M. Nicolas BURQ	Examinateur	Université Paris-Sud
M. Jean-Marc DELORT	Examinateur	Université Paris 13
M. Daniel TATARU	Examinateur	University of California, Berkeley
M. Yvan MARTEL	Directeur de thèse	École polytechnique
M. Frank MERLE	Directeur de thèse	Université de Cergy-Pontoise & IHES

Sur la dynamique d'équations des ondes avec non-linéarité énergie critique focalisante

Résumé. Cette thèse est consacrée à l'étude du comportement global des solutions de l'équation des ondes énergie-critique focalisante. On s'intéresse tout spécialement à la description de la dynamique du système dans l'espace d'énergie. Nous développons une variante de la méthode d'énergie qui permet de construire des solutions explosives de type II, instables. Ensuite, par une démarche similaire, nous donnons le premier exemple d'une solution radiale de l'équation des ondes énergie-critique qui converge dans l'espace d'énergie vers une superposition de deux états stationnaires (bulles). En appliquant notre méthode au cas de l'équation des ondes des applications harmoniques (wave map), nous obtenons des solutions de type bulle-antibulle, en toute classe d'équivariance $k > 2$. Pour l'équation des ondes énergie-critique radiale, nous étudions également le lien entre la vitesse de l'explosion de type II et la limite faible de la solution au moment de l'explosion. Finalement, nous montrons qu'il est impossible qu'une solution radiale converge vers une superposition de deux bulles ayant les signes opposés.

Mots-clés : Équation des ondes, non linéarité énergie-critique, explosion, multi-soliton

---

On the dynamics of energy-critical focusing wave equations

Abstract. In this thesis we study the global behavior of solutions of the energy-critical focusing nonlinear wave equation, with a special emphasis on the description of the dynamics in the energy space. We develop a new approach, based on the energy method, to constructing unstable type II blow-up solutions. Next, we give the first example of a radial two-bubble solution of the energy-critical wave equation. By implementing this construction in the case of the equivariant wave map equation, we obtain bubble-antibubble solutions in equivariance classes $k > 2$. We also study the relationship between the speed of a type II blow-up and the weak limit of the solution at the blow-up time. Finally, we prove that there are no pure radial two-bubbles with opposite signs for the energy-critical wave equation.

Keywords. Wave equation, energy-critical nonlinearity, blow-up, multisoliton

Remerciements

Je remercie mes directeurs de thèse Yvan Martel et Frank Merle. Yvan m'a introduit dans son domaine de recherche avec beaucoup de patience et de bienveillance dès mon stage de Master. Les longues discussions avec lui ont été non seulement essentielles pour l'avancement de la thèse, mais aussi très agréables. J'ai également beaucoup apprécié les bons conseils de Frank, ses critiques, son point de vue.

Je remercie Fabrice Planchon et Wilhelm Schlag d'avoir accepté d'être rapporteurs de cette thèse et d'avoir lu le manuscrit dans de brefs délais.

Je remercie Nicolas Burq, Jean-Marc Delort, Carlos Kenig et Daniel Tataru de m'avoir fait l'honneur de constituer le jury de la soutenance.

Je remercie Pierre Raphaël. Grâce à son cours de M1 je me suis intéressé à l'analyse des EDP. C'est aussi lui qui m'a présenté à Yvan.

Je remercie Grégoire Allaire, qui m'a beaucoup aidé en tant que responsable de mon parcours de M2. Son soutien lors du processus de recrutement des doctorants a également été exrêmement précieux. Je remercie aussi Ugo Boscain et François Golse d'avoir soutenu mes différentes candidatures.

Je remercie l'Université de Chicago et le MSRI à Berkeley. J'ai eu la chance d'y passer quelques semaines pendant ma thèse. Je remercie également le Département de Mathématiques de l'Université de Chicago d'avoir soutenu ma candidature pour un stage post-doctoral.

Je remercie Carole, Marine et Pascale de m'avoir aidé avec les différents soucis administratifs que j'ai pu avoir. Elles savent que je ne suis pas doué pour cela. Je remercie également l'administration de l'EDX et l'équipe des informaticiens : Dahn, David et Jean-Luc.

Je remercie les membres et les anciens membres du CMLS grâce auxquels il a été si agréable de venir au labo et de se retrouver après le déjeuner pour une tasse du délicieux café social, ou bien pour le traditionnel goûter du lundi. Je remercie les doctorants et les anciens doctorants du CMLS, surtout mon fidèle co-bureau Thomas et mes amis Tatiana, Sandrine, Matilde, Yakine, Timofey et Benoît.

Je remercie Joseph, qui a eu la gentillesse de lire le chapitre introductif de cette thèse et de corriger, avec beaucoup de soin, de nombreuses fautes linguistiques.

Je remercie l'équipe pédagogique du Département de Mesures Physiques de l'IUT d'Orsay, où j'ai été moniteur pendant deux ans, notamment Maximilien Lagron, Sabine Marduel et Thierry Guilbaud, ainsi que mes étudiants, les victimes de cette expérience.

Je remercie à cette occasion mes professeurs qui ont affuté mon goût pour les mathématiques pendant ces dix dernières années : Józef Tymicki et Beata Niewczas (au lycée "Unii Lubelskiej" à Lublin) ; Agnieszka Bojanowska-Jackowska, Marcin Kuczma, Stanisław Kwapień, Krzysztof Oleszkiewicz, Waldemar Pałuba, Henryk Toruńczyk et Jarosław Wiśniewski (à l'Université de Varsovie) ; Davide Barilari, Jean-Yves Chemin, Patrick Gérard, Radu Ignat, Constantin Pankrashkin, Geneviève Raugel et Filippo Santambrogio (en études de Master).

Je remercie mes parents. Je peux toujours compter sur eux. Je n'imagine pas non plus la vie sans le soutien moral de mes soeurs Agnieszka et Ewa, ou l'humour discret de mes grand-parents.

Je remercie tous mes amis que je n'ai pas encore mentionnés, pour les bons moments passés ensemble et les conversations passionnantes, en particulier Michał K., Michał Z., Basia, Piotr C., Julianna, Paweł, Adela, Piotr K., Kasia, Michał M., Łukasz, Charles, Annalaura, Thibault, Brune, Kamila, Robert et l'équipe de MEASE.

Je remercie enfin tous ceux qui ont préparé avec moi le pot de thèse, ainsi que Mme Elżbieta Wójcik du restaurant "Koncertowa" à Lublin pour ses conseils culinaires.

Contents

Chapitre I

Introduction (version française)

1 Généralités sur des équations des ondes

L'objet de ce mémoire est d'étudier certains phénomènes non linéaires dans le comportement dynamique de solutions d'équations des ondes. On considère des solutions radiales de l'équation des ondes avec la non-linéarité focalisante énergie critique en dimension $N \geq 3$:

$$\partial_t^2 u(t,x) = \Delta u(t,x) + |u(t,x)|^{\frac{4}{N-2}} u(t,x), \qquad (t,x) \in \mathbb{R} \times \mathbb{R}^N, \tag{NLW}$$

et des solutions k-équivariantes de l'équation des ondes des applications harmoniques (wave map) de $\mathbb{R}^{1+2}$ dans la sphère $\mathbb{S}^2$:

$$\partial_t^2 u(t,r) = \partial_r^2 u(t,r) + \frac{1}{r}\partial_r u(t,r) - \frac{k^2}{2r^2}\sin(2u(t,r)), \qquad (t,r) \in \mathbb{R} \times (0,+\infty). \tag{WM}$$

Avant de commenter les principaux résultats de la thèse, plaçons brièvement le sujet de la théorie globale des équations des ondes dans une perspective historique.

1.1 Préliminaires

Les équations des ondes non linéaires les plus simples s'écrivent sous la forme

$$\begin{cases} \partial_t^2 u(t,x) = \Delta u(t,x) + f(x, u(t,x)), & (t,x) \in \mathbb{R} \times \mathbb{R}^N, \\ u(t_0,x) = u_0(x),\ \partial_t u(t_0,x) = \dot u_0(x), & x \in \mathbb{R}^N, \end{cases} \tag{1.1}$$

où Δ est le laplacien en N variables spatiales et f est une fonction scalaire suffisamment régulière de 2 variables réelles telle que $f(x,0) = 0$. Cette équation possède une structure hamiltonienne naturelle. Pour le voir, posons $F(x,u) := \int_0^u f(x,v)\,\mathrm{d}v$ et définissons la *fonctionnelle d'énergie* :

$$E : C_0^\infty(\mathbb{R}^N) \times C_0^\infty(\mathbb{R}^N) \ni \boldsymbol{u} = (u, \dot u) \mapsto E(\boldsymbol{u}) \in \mathbb{R},$$
$$E(\boldsymbol{u}) := \int_{\mathbb{R}^N} \frac{1}{2}|\dot u(x)|^2 + \frac{1}{2}|\nabla u(x)|^2 - F(x, u(x))\,\mathrm{d}x.$$

L'équation (1.1) peut être écrite de manière équivalente sous la forme

$$\begin{cases} \partial_t \boldsymbol{u} = J \circ \mathrm{D}E(\boldsymbol{u}), \\ \boldsymbol{u}(t_0) = \boldsymbol{u}_0, \end{cases} \tag{1.2}$$

où D est la dérivée de Fréchet, $J := \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix}$ représente la forme symplectique naturelle sur $C_0^\infty \times C_0^\infty$ et $\boldsymbol{u} = (u, \dot u)$ est un élément de l'espace des phases $C_0^\infty \times C_0^\infty$.

On note $\langle \boldsymbol{u}, \boldsymbol{v}\rangle := \int_{\mathbb{R}^N} u\cdot v + \dot u\cdot \dot v\,\mathrm{d}x$. Soit $\boldsymbol{u}(t)$ une solution classique de l'équation (1.2) sur l'intervalle (t_1, t_2). D'après la règle de dérivation d'une fonction composée on a

$$\frac{\mathrm{d}}{\mathrm{d}t}E(\boldsymbol{u}(t)) = \langle \mathrm{D}E(\boldsymbol{u}(t)), \partial_t \boldsymbol{u}(t)\rangle = \langle \mathrm{D}E(\boldsymbol{u}(t)), J\circ \mathrm{D}E(\boldsymbol{u}(t))\rangle = 0, \tag{1.3}$$

donc $E(\boldsymbol{u}(t))$ est une loi de conservation.

1.2 Remarques historiques

L'histoire de la résolution locale en temps du problème de Cauchy remonte au moins jusqu'au théorème de Cauchy-Kowalevski, qui garantit, dans le cas d'une non-linéarité analytique, l'existence d'une unique solution analytique pour toute donnée initiale analytique. Cependant, la résolution globale en temps est essentielle du point de vue physique, et le premier résultat dans cette direction a été obtenu par Jörgens [43], suivi par des versions plus abstraites de Browder [7] et Segal [83]. Énonçons le résultat dans le cas le plus simple de l'équation de Klein-Gordon cubique, défocalisante, en dimension $1+3$, autrement dit l'équation (1.1) avec $f(x,u) = -u-u^3$:

$$\partial_t^2 u(t,x) = \Delta u(t,x) - u(t,x) - u(t,x)^3, \qquad (t,x)\in\mathbb{R}\times\mathbb{R}^3. \tag{1.4}$$

Théorème 1.1 (Jörgens, Browder, Segal)**.** *Pour toute donnée initiale $\boldsymbol{u}(t_0) = \boldsymbol{u}_0 \in H^1(\mathbb{R}^3)\times L^2(\mathbb{R}^3)$, il existe une unique solution globale $\boldsymbol{u}(t)$ de* (1.4)*. De plus, si $\boldsymbol{u}_0 \in H^{k+1}\times H^k$ avec $k\in\mathbb{N}$, alors $\boldsymbol{u}\in C(\mathbb{R}; H^{k+1}\times H^k)$.*

Remarque 1.2. Plus loin dans cette section nous clarifions la notion de solution dans le cadre non lisse.

La partie essentielle de la preuve consiste à établir le résultat suivant d'existence locale :

Proposition 1.3. *Pour toute donnée initiale $\boldsymbol{u}(t_0) = \boldsymbol{u}_0\in H^1(\mathbb{R}^3)\times L^2(\mathbb{R}^3)$, il existe une unique solution maximale $\boldsymbol{u}(t) : (T_-, T_+)\to H^1\times L^2$ de* (1.4)*. De plus, si $\boldsymbol{u}_0\in H^{k+1}\times H^k$ avec $k\in\mathbb{N}$, alors $\boldsymbol{u}\in C((T_-,T_+); H^{k+1}\times H^k)$. Si $T_+ < +\infty$, alors $\lim_{t\to T_+}\|\boldsymbol{u}(t)\|_{H^1\times L^2} = +\infty$ (de même pour $T_- > -\infty$). Enfin, sur les intervalles de temps compacts, la solution dépend continûment de la donnée initiale dans la topologie $H^{k+1}\times H^k$ pour tout $k\in\mathbb{N}$.*

Par (1.3) et l'argument habituel d'approximation, on voit que si $\boldsymbol{u}(t)$ est la solution donnée par la Proposition 1.3, alors l'énergie

$$E(\boldsymbol{u}(t)) = \int_{\mathbb{R}^N}\frac12|\dot u(t,x)|^2 + \frac12|\nabla u(t,x)|^2 + \frac12|u(t,x)|^2 + \frac14|u(t,x)|^4\,\mathrm{d}x$$

a une valeur constante. Mais cela implique que la norme $\|\boldsymbol{u}(t)\|_{H^1\times L^2}$ reste bornée, ce qui démontre le Théorème 1.1.

Comme la non-linéarité est défocalisante, on peut conjecturer que les solutions doivent décroître ponctuellement au moins comme les solutions de l'équation linéaire, rendant ainsi les effets non linéaires négligeables. Par conséquent, le comportement asymptotique de toute solution serait celui d'une solution de l'équation de Klein-Gordon libre. Ce problème s'est avéré difficile et a été résolu dans les années 80 par Brenner [6], avec des contributions importantes de Morawetz et Strauss [70].

Théorème 1.4 (Brenner, Morawetz, Strauss)**.** *Soit $\boldsymbol{u}(t)$ une solution de* (1.4) *et soit $U_0(t)$ le propagateur de Klein-Gordon linéaire. Alors $U_0(-t)\boldsymbol{u}(t)$ a une limite forte dans l'espace $H^1\times L^2$ quand $t\to\pm\infty$.*

Remarque 1.5. L'exsitence de la limite faible est relativement élémentaire et est connue depuis les années 60.

Pour ainsi dire, le comportement dynamique des solutions de l'équation (1.4) n'est pas très différent de la dynamique de l'équation linéaire, au moins dans l'espace d'énergie. Dans la justification de ce résultat, le caractère défocalisant de la non-linéarité ainsi que sa croissance en $\pm\infty$ (u^3) jouent un rôle décisif. Il est naturel d'examiner le cas d'une non-linéarité focalisante ou avec une croissance plus rapide en $\pm\infty$ (par exemple u^5), et c'est précisément ce qui va nous occuper dans les deux prochains paragraphes.

1.3 Effets de la non-linéarité focalisante

Considérons l'équation (1.1) avec $f(x,u) = -u+u^3$ (dite l'équation de Klein-Gordon cubique focalisante), en dimension $N = 3$:

$$\partial_t^2 u(t,x) = \Delta u(t,x) - u(t,x) + u(t,x)^3, \qquad (t,x) \in \mathbb{R}\times\mathbb{R}^3. \tag{1.5}$$

Quant à le théorie de Cauchy locale, rien ne change et la Proposition 1.3 reste vraie. Pourtant, la fonctionnelle d'énergie dans le cas focalisant s'écrit :

$$E(\boldsymbol{u}) = \int_{\mathbb{R}^3} \frac{1}{2}|\dot u|^2 + \frac{1}{2}|\nabla u|^2 + \frac{1}{2}|u|^2 - \frac{1}{4}|u|^4 \,\mathrm{d}x,$$

donc, à cause du "mauvais" signe devant le terme $|u|^4$, elle ne permet plus de contrôler pour tout temps la norme $H^1\times L^2$. Keller [45] a été le premier à observer que le Théorème 1.1 n'est plus valable dans le cas focalisant. Une manière simple de s'en convaincre est d'analyser les solutions constantes en espace, donc les solutions de l'équation ordinaire $u''(t) = -u(t)+u(t)^3$. On voit que la solution associée à une donnée initiale constante en espace et suffisamment grande devient $+\infty$ en temps fini. En utilisant la propriété de la vitesse finie de propagation, on peut construire une donnée initiale dans C_0^∞, pour laquelle la solution forme une singularité en temps fini.

Une autre conséquence de la non-linéarité focalisante est l'existence d'états stationnaires W, qui sont les solutions du problème elliptique

$$\Delta W(x) - W(x) + W(x)^3 = 0, \qquad x\in\mathbb{R}^3. \tag{1.6}$$

La donnée initiale $\boldsymbol{u}_0 := \boldsymbol{W} := (W,0)$ conduit à une solution de (1.5) constante en temps. Parmi toutes les solutions stationnaires, *l'état fondamental* est particulièrement important. Il peut être caractérisé comme l'unique (à une translation près) solution positive de (1.6), ou comme la solution qui minimise l'énergie potentielle $E(u) := \int_{\mathbb{R}^3}\frac{1}{2}|\nabla u|^2 + \frac{1}{2}|u|^2 - \frac{1}{4}|u|^4\,\mathrm{d}x$. Dans le langage du calcul des variations, $E(u)$ crée un puits de potentiel (appelé aussi *cuvette*) pour u suffisamment petit dans l'espace d'énergie, et W est le point col. Payne et Sattinger [75] ont montré que cette structure variationnelle est liée au rôle de $E(W)$ comme *seuil d'énergie* pour les différents types du comportement dynamique : si une solution d'énergie $E(\boldsymbol{u}_0) < E(W)$ est à l'intérieur de la cuvette ($\|u_0\|_{H^1} \leq \|W\|_{H^1}$), elle y reste pour toujours, et si elle est à l'extérieur ($\|u_0\|_{H^1} > \|W\|_{H^1}$), elle subit une explosion en temps fini.

Puisque les solutions dans le puits de potentiel sont globales, il est naturel d'étudier si elles ont le comportement linéaire quand $t\to+\infty$, comme dans le cas défocalisant. Une approche générale de ce problème, appelé *Conjecture de l'état fondamental* ou *Conjecture du seuil*, a été développée par Kenig et Merle [46, 47] pour l'équation des ondes énergie critique et pour l'équation de Schrödinger énergie critique radiale. Dans le cas de l'équation de Klein-Gordon cubique, le preuve est due à Ibrahim, Masmoudi et Nakanishi [38].

Théorème 1.6 (Ibrahim, Masmoudi, Nakanishi). *Soit $\boldsymbol{u}(t)$ une solution de* (1.5) *telle que $E(\boldsymbol{u}_0) < E(W)$ et $\|\boldsymbol{u}_0\|_{H^1\times L^2} \leq \|W\|_{H^1}$. Alors $U_0(-t)\boldsymbol{u}(t)$ a une limite forte dans l'espace $H^1 \times L^2$ quand $t \to \pm\infty$.*

Après avoir classifié les solutions en-dessous du seuil d'énergie $E(W)$, l'objectif suivant devrait être de comprendre le comportement dynamique *légèrement au-dessus* de ce seuil. Une solution particulièrement élégante à ce problème, exhibant un lien avec la théorie des variétés invariantes, a été proposée par Nakanishi et Schlag [71, 72].

1.4 Non-linéarités énergie critiques

Revenons pour un instant à l'équation générale (1.1). On a vu que la structure du problème permet de contrôler, peut-être de manière indirecte comme dans le cas focalisant, la norme homogène d'énergie

$$\|\boldsymbol{u}(t)\|^2_{\dot H^1\times L^2} := \int_{\mathbb{R}^N} |\partial_t u(t,x)|^2 + |\nabla u(t,x)|^2\,\mathrm{d}x,$$

ce qui nous donne une certaine information sur la régularité de la solution (dans le cas de Klein-Gordon on peut contrôler également $\|u\|_{L^2}$, mais c'est moins intéressant du point de vue de la régularité). En fonction de la croissance de $|f(x,u)|$ quand $|u| \to +\infty$, cette information peut être ou ne pas être suffisante pour exclure l'explosion de la solution. On va étudier maintenant de manière heuristique la possibilité de formation d'une singularité telle que la norme d'énergie reste bornée. Une singularité de ce type est appelée *explosion de type II* ou *explosion géométrique*, le terme *explosion de type I* désignant quant à lui la situation où la norme d'énergie tend vers $+\infty$ en temps fini.

Pour $\boldsymbol{u}(x) = (u(x), \dot u(x))$ et $\lambda > 0$, on définit le changement d'échelle énergie critique

$$u_\lambda(x) := \frac{1}{\lambda^{\frac{N-2}{2}}}u\Big(\frac{x}{\lambda}\Big), \quad \boldsymbol{u}_\lambda(x) := \Big(\frac{1}{\lambda^{\frac{N-2}{2}}}u\Big(\frac{x}{\lambda}\Big), \frac{1}{\lambda^{\frac{N}{2}}}\dot u\Big(\frac{x}{\lambda}\Big)\Big)$$

Une intégration par changement de variables montre que $\|\boldsymbol{u}_\lambda\|_{\dot H^1\times L^2} = \|\boldsymbol{u}\|_{\dot H^1\times L^2}$. Pourtant, la norme d'énergie *se concentre près de* $x = 0$ quand $\lambda \to 0$. Supposons pour simplifier que $f(x,u) = f(u)$ ne dépend pas de x et que $N \geq 3$. Si $|f(u)| \lesssim |u|^p$ quand $|u| \to +\infty$, on obtient

$$|f(u_\lambda)| \lesssim \frac{1}{\lambda^{\frac{p(N-2)}{2}-\frac{N-2}{2}}}\big|\big(|u|^p\big)_\lambda\big|, \quad \text{pour } \lambda \text{ petit},$$

alors que

$$\Delta u_\lambda = \frac{1}{\lambda^2}\big(\Delta u\big)_\lambda.$$

Si $\frac{p(N-2)}{2} - \frac{N-2}{2} < 2 \Leftrightarrow p < \frac{N+2}{N-2}$, on voit que l'effet de la non-linéarité est négligeable par rapport au laplacien quand $\lambda \to 0$. Intuitivement, la concentration de la norme d'énergie rend négligeables les effets non linéaires, en particulier ne peut pas conduire à la formation d'une singularité. Par conséquent, la seule manière de former une singularité est l'explosion de la norme d'énergie, ce qui fait partie de l'énoncé de la Proposition 1.3. On appelle ce type de non-linéarité *énergie sous-critique*. En revanche, si $p > \frac{N+2}{N-2}$, alors la non-linéarité constitue la partie dominante quand $\lambda \to 0$. Cela signifie qu'une concentration d'une quantité arbitrairement petite de la norme énergétique pourrait potentiellement rendre les effets non linéaires décisifs, conduisant ainsi à une formation d'une singularité. On appelle une telle non-linéarité *énergie sur-critique*. Le cas limite, $p = \frac{N+2}{N-2}$, est dit *énergie critique*. Dans cette dernière situation, la non-linéarité agit avec la même force que le laplacien, a priori formant une singularité de type II en temps fini.

L'équation des ondes énergie critique la plus simple est probablement l'équation semi-linéaire avec la non-linéarité défocalisante de type puissance :

$$\partial_t^2 u(t,x) = \Delta u(t,x) - |u(t,x)|^{\frac{4}{N-2}} u(t,x), \qquad (t,x) \in \mathbb{R} \times \mathbb{R}^N. \tag{1.7}$$

Le caractère bien posé dans l'espace d'énergie $\dot H^1 \times L^2$ a été démontré par Shatah et Struwe [84], avec des contributions importantes de Kapitanski [44], et Ginibre, Soffer et Velo [32]. Observons que pour tout $\boldsymbol{u} = (u, \dot u) \in \dot H^1 \times L^2$, l'énergie $E(\boldsymbol{u}) = \int_{\mathbb{R}^N} \frac{1}{2}|\dot u|^2 + \frac{1}{2}|\nabla u|^2 + \frac{N-2}{2N}|u|^{\frac{2N}{N-2}}\,\mathrm{d}x$ est bien définie grâce à l'injection de Sobolev critique. De plus, on vérifie facilement que $E(\boldsymbol{u}_\lambda) = E(\boldsymbol{u})$. Enfin, l'équation est invariante par le changement d'échelle énergie critique : si $(u(t,x), \dot u(t,x))$ est une solution de (1.7) sur un intervalle $I \ni t_0$, alors pour tout $\lambda > 0$ on peut former une autre solution $\big(u\big(t_0 + \frac{t}{\lambda}, \frac{x}{\lambda}\big)\big)$, définie sur un intervalle de temps qui contient $t = 0$.

Il s'avère que pour les solutions de (1.7), la norme d'énergie ne peut pas se concentrer, et qu'un équivalent du Théorème 1.1 est vrai, comme l'ont démontré Struwe [88] dans le cas radial et Grillakis [34] sans l'hypothèse de symétrie. En outre, les solutions convergent vers des ondes linéaires quand $t \to \pm\infty$. Dans le livre [91, Chapter 5] on peut trouver les détails historiques concernant la résolution de ce problème.

Dans cette thèse nous nous concentrerons principalement sur l'équation (NLW), l'analogue focalisant de (1.7). L'énergie est donnée par

$$E(\boldsymbol{u}) = \int_{\mathbb{R}^N} \frac{1}{2}|\dot u|^2 + \frac{1}{2}|\nabla u|^2 - \frac{N-2}{2N}|u|^{\frac{2N}{N-2}}\,\mathrm{d}x.$$

Le caractère bien posé se démontre en modifiant les preuves pour le cas défocalisant, cf. [46]. On obtient le résultat suivant.

Proposition 1.7. *Pour toute donnée initiale $\boldsymbol{u}_0 \in \dot H^1(\mathbb{R}^N) \times L^2(\mathbb{R}^N)$, il existe une unique solution maximale $\boldsymbol{u}(t) : (T_-, T_+) \to \dot H^1 \times L^2$ de* (NLW). *Si $\|\boldsymbol{u}_0\|_{\dot H^1 \times L^2}$ est suffisamment petit, alors $T_- = -\infty$, $T_+ = +\infty$ et $\boldsymbol{u}(t)$ converge vers une onde linéaire dans les deux directions du temps. La solution est continue par rapport à la donnée initiale dans la topologie $\dot H^1 \times L^2$.*

Il faut noter une différence majeure par rapport au cas sous-critique traité ci-dessus. Si $T_+ < +\infty$, il n'est plus garanti que $\lim_{t\to T_+} \|\boldsymbol{u}\|_{\dot H^1 \times L^2} = +\infty$, ce dont la raison heuristique a été présentée au début de ce paragraphe. On a tout de même le critère abstrait d'explosion : si $T_+ < +\infty$ et K est un sous-ensemble compact de $\dot H^1 \times L^2$, alors il existe $\tau > 0$ tel que $\boldsymbol{u}(t) \notin K$ pour $t \in [T_+ - \tau, T_+)$.

Le problème elliptique

$$\Delta W(x) + |W(x)|^{\frac{4}{N-2}} W(x) = 0, \qquad x \in \mathbb{R}^N \tag{1.8}$$

(appelé l'équation de Yamabe) possède la solution positive explicite suivante :

$$W(x) = \Big(1 + \frac{|x|^2}{N(N-2)}\Big)^{-\frac{N-2}{2}}.$$

En raison de l'invariance de l'équation par rapport au changement d'échelle énergie-critique, pour tout $\lambda > 0$, la fonction $W_\lambda(x)$ est également une solution de (1.8). Il est connu que ce sont toutes les solutions à symétrie radiale d'énergie finie et, à une translation près, toutes les solutions positives de (1.8). Cependant, le problème non radial est difficile, cf. [73]. On obtient les solutions stationnaires à symétrie radiale de (NLW) $\boldsymbol{W}_\lambda = (W_\lambda, 0)$. Le rôle de W_λ comme le point col pour l'énergie potentielle $E(u) = \int \frac{1}{2}|\nabla u|^2 - \frac{N-2}{2N}|u|^{\frac{2N}{N-2}}\,\mathrm{d}x$ résulte des travaux d'Aubin [1] et de Talenti [90].

La Conjecture de l'état fondamental pour l'équation des ondes focalisante énergie-critique a été démontrée dans les travaux de Kenig et Merle [46] pour $N \in \{3,4,5\}$. Pour $N = 3$, Krieger, Schlag et Tataru [53] ont construit des solutions qui développent une singularité en temps fini, et qui restent pendant tout le temps d'existence dans un voisinage arbitrairement petit (dans l'espace d'énergie) de la famille des états fondamentaux $\boldsymbol{W}_\lambda$. En particulier, cette construction a donné le premier exemple d'une explosion de type II. Duyckaerts, Kenig et Merle [23, 25] ont montré que toute solution de (NLW) qui explose dans un voisinage de l'ensemble $\{\boldsymbol{W}_\lambda\}$, se décompose asymptotiquement dans l'espace d'énergie en une somme de l'état fondamental, après un changement d'échelle et une transformée de Lorentz, et d'un profil asymptotique. Plus précisément, dans le cas d'une symétrie radiale, si $T_+ < +\infty$, alors il existe une fonction positive $\lambda(t) \ll T_+ - t$ et $\boldsymbol{u}_0^* \in \dot{H}^1 \times L^2$ tels que

$$\lim_{t\to T_+} \|\boldsymbol{u}(t) - \boldsymbol{u}_0^* - \boldsymbol{W}_{\lambda(t)}\|_{\dot{H}^1\times L^2} = 0.$$

En dimension $N = 3$, pour les solutions à symétrie radiale, les mêmes auteurs [26] ont donné une classification complète du comportement dynamique des solutions de (NLW) dans l'espace d'énergie. Ils ont montré que si $T_+ = +\infty$, alors il existe $J \in \mathbb{N}$, des fonctions strictement positives $\lambda_1(t) \ll \lambda_2(t) \ll \ldots \lambda_J(t) \ll t$ et une solution $\boldsymbol{u}^*(t)$ de l'équation des ondes linéaire tels que

$$\lim_{t\to+\infty} \Big\|\boldsymbol{u}(t) - \boldsymbol{u}^*(t) - \sum_{j=1}^{J} \pm\boldsymbol{W}_{\lambda_j(t)}\Big\|_{\dot{H}^1\times L^2} = 0. \tag{1.9}$$

Si $T_+ < +\infty$, alors soit $\lim_{t\to T_+} \|\boldsymbol{u}(t)\|_{\dot{H}^1\times L^2} = +\infty$, soit il existe $J \in \mathbb{N}$, des fonctions strictement positives $\lambda_1(t) \ll \lambda_2(t) \ll \ldots \lambda_J(t) \ll T_+ - t$ et $\boldsymbol{u}_0^* \in \dot{H} \times L^2$ tels que

$$\lim_{t\to T_+} \Big\|\boldsymbol{u}(t) - \boldsymbol{u}_0^* - \sum_{j=1}^{J} \pm\boldsymbol{W}_{\lambda_j(t)}\Big\|_{\dot{H}^1\times L^2} = 0. \tag{1.10}$$

Krieger, Nakanishi et Schlag [49] ont étudié la dynamique dans un voisinage des états fondamentaux d'un point de vue différent, en relation avec la théorie des variétés invariantes. En particulier, ils ont construit la *variété centre-stable*, qui est une hypersurface de classe C^1 qui contient $\{\boldsymbol{W}_\lambda\}$, et ils ont montré que cette hypersurface sépare l'ensemble des solutions ayant le comportement asymptotiquement linéaire pour les temps positifs de l'ensemble des solutions qui développent une singularité de type I. Les solutions sur la variété sont celles qui restent dans un voisinage de $\{\boldsymbol{W}_\lambda\}$ jusqu'à la fin de leur temps d'existence.

Remarque 1.8. Une théorie globale des équations des ondes super-critiques pour de grandes données initiales, même dans le cas défocalisant, semble actuellement inaccessible. À la connaissance de l'auteur, on dispose uniquement de résultats conditionnels sur la dynamique des solutions (aussi bien dans le cas focalisant que défocalisant), cf. Kenig et Merle [48]. Dans le cas où la non-linéarité défocalisante est une puissance, Krieger et Schlag [51] ont montré l'existence de grandes données initiales telles que la solution existe pour tout temps.

1.5 Autres modèles critiques

Historiquement, le modèle énergie critique le plus étudié est *le flot de la chaleur des applications harmoniques* entre deux surfaces, surtout entre deux sphères de dimension 2. Pour l'état de l'art dans ce domaine, le lecteur peut consulter Topping [94], ainsi que les références citées dans cet article. En particulier, c'est dans ce cas-là que les premiers résultats de classification du comportement à l'explosion (*bubbling*) ont été obtenus.

Les termes “une explosion de type I” et “une explosion de type II” proviennent des travaux sur l'équation de la chaleur sur-critique $\partial_t u = \Delta u + |u|^p u$, cf. [63].

L'équation de Schrödinger masse critique (ou L^2 critique) a été étudiée dans de nombreux travaux, aussi en vue des applications physiques. Sous certains aspects, elle ressemble à l'équation de Korteweg-de Vries généralisée L^2 critique. L'article de revue [61] présente les avancées majeures récentes dans la compréhension de ces deux modèles.

Parmi les équations hamiltoniennes énergie critique, notons l'équation de Schrödinger semi-linéaire : $i\partial_t u + \Delta u + |u|^{\frac{4}{N-2}} u = 0$ sur $\mathbb{R}^N$, et l'équation de Schrödinger des applications harmoniques (Schrödinger maps) de $\mathbb{R}^2$ dans la sphère $\mathbb{S}^2 \subset \mathbb{R}^3$: $\partial_t u = u \wedge \Delta u$, où $\wedge$ est le produit vectoriel dans $\mathbb{R}^3$. Par la suite, on évoquera parfois ces deux modèles, par souci de comparaison avec les équations des ondes.

Remarque 1.9. Les solutions des équations énergie critiques ont souvent un comportement dynamique compliqué, mais d'un autre côté elles possèdent une structure supplémentaire d'invariance par rapport au changement d'échelle énergie critique, ce qui est essentiel pour beaucoup de résultats de classification.

1.6 Éléments de la théorie de Cauchy

Dans ce paragraphe on donne quelques commentaires sur la notion de solution et de caractère bien posé. Il sera commode de noter $X^s := \dot H^{s+1} \cap \dot H^1$ pour $s \geq 0$.

Définition 1.10. Soit $t_0 \in (t_1, t_2) \subset \mathbb{R}$, $\boldsymbol{u}_0 = (u_0, \dot u_0) \in X^s \times H^s$ et $h \in L^1((t_1, t_2); H^s)$. Pour $t \in (t_1, t_2)$ posons

$$\begin{aligned} u(t) &:= \cos\big((t-t_0)\sqrt{-\Delta}\big)u_0 + \frac{\sin\big((t-t_0)\sqrt{-\Delta}\big)}{\sqrt{-\Delta}}\dot u_0 \\ &\quad + \int_{t_0}^t \frac{\sin\big((t-s)\sqrt{-\Delta}\big)}{\sqrt{-\Delta}} h(s)\,\mathrm{d}s, \\ \dot u(t) &:= -\sin\big((t-t_0)\sqrt{-\Delta}\big)\sqrt{-\Delta}u_0 + \cos\big((t-t_0)\sqrt{-\Delta}\big)\dot u_0 \\ &\quad + \int_{t_0}^t \cos\big((t-s)\sqrt{-\Delta}\big)h(s)\,\mathrm{d}s. \end{aligned}$$

On appelle $\boldsymbol{u}(t) := (u(t), \dot u(t))$ *la solution* du problème

$$\begin{cases} \partial_t^2 u = \Delta u + h, \\ (u, \partial_t u)_{t=t_0} = (u_0, \dot u_0). \end{cases} \tag{1.11}$$

Notons qu'il s'ensuit directement que $\boldsymbol{u} \in C((t_1,t_2), X^s \times H^s)$ (on dit souvent que “l'équation des ondes fait gagner une dérivée”). Une version quantitative de ce fait est donnée par l'inégalité fondamentale suivante :

$$\|\boldsymbol{u}(t)\|_{X^s\times H^s} \leq \|\boldsymbol{u}_0\|_{X^s\times H^s} + \Big|\int_{t_0}^t \|h(s)\|_{H^s}\,\mathrm{d}s\Big|, \tag{1.12}$$

que l'on appelle *l'estimée d'énergie*. L'existence et l'unicité des solutions faibles du problème linéaire (1.11), ainsi que l'estimée d'énergie (dans un cadre plus général des systèmes symétriques) ont été démontrées par Friedrichs [30].

Définition 1.11. Supposons que $\partial_u f(x, 0) = 0$. On dit que l'équation (1.1) est *localement bien posée* dans $X^s \times H^s$ si

- pour tout $\boldsymbol{u}_0 \in X^s \times H^s$ et $t_0 \in \mathbb{R}$ il existe un intervalle du temps $(t_1, t_2) \ni t_0$ et une unique fonction $\boldsymbol{u}(t) = (u(t), \dot{u}(t))$ telle que

$$f(x, u(t,x)) \in L^1((t_1, t_2); H^s) \tag{1.13}$$

  et $\boldsymbol{u}(t)$ est une solution de (1.11) avec $h(t,x) = f(x, u(t,x))$,

- $\boldsymbol{u}(t)$ dépend de $\boldsymbol{u}_0$ de façon continue dans la topologie $X^s \times H^s$.

Remarque 1.12. Pour des fonctions lisses, cette définition coïncide avec la définition habituelle d'une solution classique d'une équation différentielle.

Remarque 1.13. En général, au lieu de $-\Delta$, on devrait considérer l'opérateur $T := -\Delta - \partial_u f(x, 0)$. Sous des hypothèses naturelles sur T, on peut définir la notion de solution de manière similaire, en remplaçant partout H^s par $(1 + T^{\frac{s}{2}})^{-1}L^2$ et X^s par $(T^{\frac{1}{2}} + T^{\frac{s+1}{2}})^{-1}L^2$.

Remarque 1.14. Dans des situations typiques, la condition (1.13) est équivalente au fait que $u(t)$ appartient à un espace fonctionnel naturel, par exemple un espace de Lebesgue sur l'espace-temps. La Définition 1.11 formule un problème de point fixe. La partie principale de sa résolution consiste à trouver des bornes de $\|f(x, u(t,x))\|_{L^1((t_1,t_2);H^s)}$ en termes de $\|h(t,x)\|_{L^1((t_1,t_2);H^s)}$, où $u(t,x)$ est la solution de (1.11). Pour cela, on utilise des inégalités de type Sobolev ou Strichartz.

Remarque 1.15. Supposons que (1.1) est localement bien posée dans $X^s \times H^s$ et que $\boldsymbol{u}(t)$ est une solution avec temps maximal d'existence $T_+ < +\infty$. Alors $\|f(x, u(t,x))\|_{L^1([t_0,T_+);H^s)} = +\infty$. Si ce n'était pas le cas, alors (1.12) impliquerait que la solution $\boldsymbol{u}(t)$ est pré-compacte dans $X^s \times H^s$ quand $t \to T_+$, ce qui est impossible.

Dans la même veine, si $T_+ = +\infty$ et $\|f(x, u(t,x))\|_{L^1([t_0,+\infty);H^s)} < +\infty$, alors $\boldsymbol{u}(t)$ a le comportement asymptotiquement linéaire dans $X^s \times H^s$ quand $t \to +\infty$.

Remarque 1.16. Dans certains cas "pathologiques", comme par exemple (NLW) en grande dimension, cette définition générale n'est pas nécessairement la bonne, et il faut remplacer la condition $f(x, u(t,x)) \in L^1((t_1, t_2); H^s)$ par une autre restriction garantissant que la solution est unique et qu'elle dépend d'une manière continue de la donnée initiale.

2 Résultats

Dans cette thèse on considère l'équation (NLW) pour les données initiales à symétrie radiale, sauf dans le Chapitre 2, où on traite également de l'équation (WM).

Dans le prolongement des travaux cités ci-dessus, on étudie le système dynamique défini par l'équation (NLW), au voisinage de l'ensemble $\{\boldsymbol{W}_\lambda\}$ dans l'espace d'énergie, c'est-à-dire les solutions $\boldsymbol{u}(t)$ de (NLW) telles que

$$\inf_{\lambda>0} \|\boldsymbol{u}(t) - \boldsymbol{W}_\lambda\|_{\dot{H}^1 \times L^2} \leq \eta, \qquad \forall t,$$

où $\eta > 0$ est une petite constante. Les Chapitres 1 et 3 sont consacrés au phénomène de l'explosion de type II.

Dans les Chapitres 2 et 4, on étudie le comportement local au voisinage d'une superposition de deux bulles à des échelles différentes, c'est-à-dire les solutions $\boldsymbol{u}(t)$ de (NLW) telles que

$$\inf_{0<\lambda\leq\alpha^*\mu} \|\boldsymbol{u}(t) - (\boldsymbol{W}_\mu \pm \boldsymbol{W}_\lambda)\|_{\dot{H}^1 \times L^2} \leq \eta, \qquad \forall t,$$

où $\eta > 0$ est une petite constante. La motivation de ce travail vient bien évidemment des résultats de classification de [26] mentionnés plus haut, cf. (1.9) et (1.10).

Cette thèse est composée des quatre articles suivants :

- Chapitre 1 – Jendrej, J. Construction of type II blow-up solutions for the energy-critical wave equation in dimension 5. *Prépublication*, arXiv:1503.05024, 2015.
- Chapitre 2 – Jendrej, J. Construction of two-bubble solutions for energy-critical wave equations. *Prépublication*, arXiv:1602.06524, 2016.
- Chapitre 3 – Jendrej, J. Bounds on the speed of type II blow-up for the energy critical wave equation in the radial case. *Int. Math. Res. Not.*, doi : 10.1093/imrn/rnv365, 2015.
- Chapitre 4 – Jendrej, J. Nonexistence of radial two-bubbles with opposite signs for the energy-critical wave equation. *Prépublication*, arXiv:1510.03965, 2015.

2.1 Construction de solutions explosives de type II

Dans cet article nous développons une nouvelle approche de la construction de solutions explosives de type II pour l'équation des ondes énergie critique. De telles solutions ont été d'abord construites pour l'équation (NLW) en dimension $N = 3$ et pour l'équation (WM) en classe d'équivariance $k = 1$ par Krieger, Schlag et Tataru [52, 53].

Ici, on considère l'équation (NLW) en dimension $N = 5$. On démontre les résultats suivants.

Théorème 2.1. *Soit* $\boldsymbol{u}_0^* = (u_0^*, \dot u_0^*) \in (\dot H^5 \cap \dot H^1) \times H^4$ *une paire de fonctions à symétrie radiale avec* $u_0^*(0) > 0$. *Il existe une solution* $\boldsymbol{u}(t)$ *de* (NLW)*, définie sur l'intervalle du temps* $(0, T_0)$*, telle que*

$$\lim_{t\to 0} \big\|\boldsymbol{u}(t) - \boldsymbol{u}_0^* - \boldsymbol{W}_{\lambda(t)}\big\|_{\dot H^1\times L^2} = 0, \tag{2.1}$$

où $\lambda(t) = \big(\frac{32}{315\pi}\big)^2\big(u^*(0,0)\big)^2 t^4$.

Théorème 2.2. *Soit* $\nu > 8$. *Il existe une solution* $\boldsymbol{u}(t)$ *de* (NLW)*, définie sur l'intervalle du temps* $(0, T_0)$*, telle que*

$$\lim_{t\to 0} \big\|\boldsymbol{u}(t) - \boldsymbol{u}_0^* - \boldsymbol{W}_{\lambda(t)}\big\|_{\dot H^1\times L^2} = 0,$$

où $\lambda(t) = t^{\nu+1}$ *et* $\boldsymbol{u}_0^*$ *est explicite.*

Le fait d'obtenir un continuum de vitesses d'explosion possibles indique que le comportement dynamique de nos solutions est fortement instable par rapport aux variations de la donnée initiale. Un peu plus précisément, on peut s'attendre à ce que l'ensemble des solutions ayant les mêmes caractéristiques dynamiques (par exemple la même vitesse d'explosion) soit de codimension infinie. On pourrait dire que l'on obtient des solutions de type Krieger-Schlag-Tataru, par opposition aux solutions stables construites par Hillairet et Raphaël [36] en dimension $N = 4$, qui peuvent exploser seulement avec des vitesses bien spécifiques. De telles solutions ont été aussi construites pour l'équation (WM) par Rodnianski et Sterbenz [80], et par Raphaël et Rodnianski [79].

Remarque 2.3. Dans ce deuxième cas, la possibilité d'une formation de singularités avait été suggérée par des expériences numériques, voir par exemple [4]. C'est l'explosion stable que l'on observe numériquement. Les auteurs de [4] suggèrent que l'ensemble des données initiales donnant lieu à une explosion est énorme. Cependant, l'existence d'un sous-ensemble ouvert de l'espace d'énergie, tel que toute donnée initiale dans ce sous-ensemble développe une singularité, reste inconnu.

Présentons maintenant les idées de la preuve du Théorème 2.1, sans nous préoccuper de tous les détails techniques.

Étape 1. Supposons que $\boldsymbol{u}(t) = (u(t), \dot u(t))$ vérifie (2.1). Soit $\boldsymbol{u}^*(t) = (u^*(t), \dot u^*(t))$ la solution de (NLW) pour la donnée initiale $\boldsymbol{u}^*(0) = \boldsymbol{u}_0^*$. Notons $\Lambda W := -\frac{\partial}{\partial\lambda} W_\lambda|_{\lambda=1}$. Par la règle de dérivation d'une fonction composée on a $\partial_t W_{\lambda(t)} = -\frac{\lambda'(t)}{\lambda(t)}\Lambda W_{\lambda(t)}$, donc on s'attend à ce que

$$(u(t), \dot u(t)) \simeq \Big(u^*(t) + W_{\lambda(t)}, \dot u^*(t) - \frac{\lambda'(t)}{\lambda(t)}\Lambda W_{\lambda(t)}\Big). \tag{2.2}$$

Il semble naturel de travailler dans l'espace d'énergie. Par un changement de variable, on obtient
$\big\|\frac{\lambda'(t)}{\lambda(t)}\Lambda W_{\lambda(t)}\big\|_{L^2} \sim \lambda'(t)$ (la solution explose en $t = 0$, d'où $\lambda'(t) > 0$). Si l'on introduit le petit paramètre $b(t) \simeq \lambda'(t)$, on peut voir (2.2) comme le début d'un développement asymptotique de $\boldsymbol{u}(t)$ dans l'espace d'énergie en puissances de $b(t)$.

Il s'avère que si

$$b'(t) = \frac{128}{105\pi} u^*(t)\sqrt{\lambda(t)}, \tag{2.3}$$

alors on peut calculer le terme suivant de ce développement et définir ainsi une *solution approchée* (que l'on appelle aussi *ansatz*)

$$\boldsymbol{\varphi}(t) := \boldsymbol{u}^*(t) + \boldsymbol{U}^{(0)}_{\lambda(t)} + b(t)\boldsymbol{U}^{(1)}_{\lambda(t)} + b^2(t)\boldsymbol{U}^{(2)}_{\lambda(t)} + b^3(t)\boldsymbol{U}^{(3)}_{\lambda(t)},$$

où $\boldsymbol{U}^{(0)} = (W, 0)$ et $\boldsymbol{U}^{(1)} = (0, -\Lambda W)$. Les définitions précises de $\boldsymbol{U}^{(2)}$ et de $\boldsymbol{U}^{(3)}$ ne sont pas essentielles. Observons que l'équation (2.3) avec $\lambda' = b$ donne la valeur $\lambda_{\text{app}}(t) = \big(\frac{32}{315\pi}\big)^2\big(u^*(0,0)\big)^2 t^4$, qui est celle de l'énoncé du théorème.

Remarque 2.4. Dans nos définitions de $\boldsymbol{U}^{(2)}$ et de $\boldsymbol{U}^{(3)}$, la décroissance en $|x| \to +\infty$ de $W(x)$ joue un rôle important, et c'est la raison pour laquelle on doit se restreindre aux dimensions $N \geq 5$. En grande dimension la non-linéarité $|u|^{\frac{4}{N-2}}u$ devient assez singulière en $u = 0$, ce qui introduit des difficultés techniques supplémentaires. On a choisi le cas le plus simple $N = 5$.

Étape 2. On considère une suite de solutions $\boldsymbol{u}_n(t)$ de (NLW) ayant comme donnée initiale

$$\boldsymbol{u}_n(t_n) = \Big(u^*(t_n) + W_{\lambda_{\text{app}}(t_n)}, \dot u^*(t_n) - \frac{\lambda'_{\text{app}}(t_n)}{\lambda_{\text{app}}(t_n)}\Lambda W(t_n)\Big),$$

où $t_n > 0$ et $\lim_{n\to+\infty} t_n = 0$ (un ajustement est à faire à cause de l'instabilité exponentielle au voisinage de $\boldsymbol{W}$, mais ce n'est pas une grave difficulté). En utilisant une condition d'orthogonalité adaptée, on décompose $\boldsymbol{u}_n(t) = \boldsymbol{\varphi}(t) + \boldsymbol{g}_n(t)$, le but étant de contrôler la taille de $\boldsymbol{g}_n(t)$, uniformément en n, sur un intervalle de temps $(t_n, t_0]$, $t_0 > 0$.

Pour cela, on introduit *une fonctionnelle mixte énergie-viriel* $H_n(t)$, qui est une petite perturbation de la fonctionnelle d'énergie $E(\boldsymbol{\varphi}(t) + \boldsymbol{g}_n(t)) - E(\boldsymbol{\varphi}(t)) - \langle \mathrm{D}E(\boldsymbol{\varphi}(t)), \boldsymbol{g}_n(t)\rangle$. Cette fonctionnelle a la propriété de coercitivité suivante :

$$\|\boldsymbol{g}_n(t)\|^2_{\dot H^1\times L^2} \lesssim H_n(t) \qquad \text{(modulo les modes instables).}$$

De plus, en utilisant le fait que $\boldsymbol{\varphi}(t)$ est un ansatz raffiné, on peut montrer qu'il existe une grande constante C_0 telle que

$$\|\boldsymbol{g}_n(t)\|_{\dot{H}^1\times L^2} \leq C_0 t^{\frac{9}{2}} \quad \Rightarrow \quad H_n'(t) \leq c_0 \cdot C_0^2 t^8,$$

avec une petite constante c_0. Un argument de continuité classique implique la borne uniforme $\|\boldsymbol{g}\|_{\dot{H}^1\times L^2} \leq C_0 t^{\frac{9}{2}}$. De cette estimation, on déduit par des techniques d'analyse des équations ordinaires que le paramètre de modulation $\lambda_n(t)$ est proche de $\lambda_{\text{app}}(t)$.

La conclusion de la deuxième étape est que

$$\|\boldsymbol{u}_n(t) - (u^*(t) + W_{\lambda_{\text{app}}}(t), \dot{u}^*(t))\|_{\dot{H}^1\times L^2} \lesssim t^3,$$

uniformément par rapport à n.

Étape 3. En utilisant la décomposition en profils de Bahouri et Gérard [3], on démontre une version de la continuité faible séquentielle du flot, ce qui permet d'obtenir notre solution $\boldsymbol{u}(t)$ comme un point d'adhérence faible de la suite $\boldsymbol{u}_n(t)$ dans l'espace d'énergie.

L'idée de construire une suite de solutions contrôlée uniformément, convergeant vers une solution singulière, a été introduite par Merle [64]. Fusionner cette technique avec la méthode d'énergie est une idée de Martel [56]. Raphaël et Szeftel [78] ont utilisé une correction par un viriel de la fonctionnelle d'énergie dans un contexte similaire, dans leur étude de solutions explosives de masse minimale pour l'équation de Schrödinger non linéaire. La première étape de notre preuve est aussi inspirée par les travaux de Martel, Merle et Raphaël [60] sur l'explosion exotique pour l'équation de Korteweg-de Vries L^2-critique. Ils ont observé que la vitesse d'explosion est reliée à la taille de l'interaction de la bulle avec le "fond", ce qui est à la base de notre construction de la solution approchée.

En réalité, la taille de cette interaction apparaît explicitement dans (2.3), ce qui sera expliqué en détail dans le Chapitre 1. Nous trouvons que le fait de mettre en lumière un lien direct entre le comportement asymptotique de $\boldsymbol{u}^*(t)$ en $x = 0$ et l'asymptotique de la fonction $\lambda(t)$ est un avantage majeur de notre méthode. Par exemple, dans le Théorème 2.2 on a $\dot{u}_0^* \equiv 0$ et $u_0^*(x) \sim |x|^{\frac{\nu-3}{2}}$ dans un voisinage de $x = 0$.

Un point délicat, que l'on n'aborde pas dans notre travail, serait de mieux comprendre la régularité des solutions construites. On ne dispose d'aucune information sur la régularité de la solution outre le fait qu'elle appartient à l'espace d'énergie. En même temps, on s'attend généralement à ce que ces solutions aient des singularités, et travailler au niveau de la régularité $\dot{H}^1 \times L^2$ permet d'éviter de les traiter directement. Observons également que, pour construire des objets aussi instables que les nôtres, il semble naturel d'utiliser des estimées d'énergie en inversant la direction du temps, à savoir dans le sens de la défocalisation de la solution. La méthode d'énergie "dans le sens de l'explosion" implique typiquement une sorte de stabilité de codimension finie.

Dans des travaux à venir, nous espérons réaliser une construction similaire pour l'équation (WM), ainsi que dans le cas de $\boldsymbol{u}_0^*$ singulier.

Remarque 2.5. À cause d'un lien avec le Chapitre 3, on signale que l'on peut aussi déterminer formellement la vitesse de l'explosion en résolvant pour $\lambda'(t)$ l'équation

$$E\Big(\Big(u^*(t) + W_{\lambda(t)}, \dot{u}^*(t) - \frac{\lambda'(t)}{\lambda(t)}\Lambda W_{\lambda(t)}\Big)\Big) = E(\boldsymbol{W}) + E(\boldsymbol{u}_0^*).$$

Remarque 2.6. Certaines notations du Chapitre 1 sont différentes de celles employées ici. Nous avons choisi d'utiliser des notations cohérentes dans toutes les sections de ce chapitre introductif.

2.2 Construction de 2-bulles

Dans le Chapitre 2 nous construisons une solution à symétrie radiale de (NLW) qui existe globalement pour les temps négatifs et qui se décompose en deux bulles d'énergie dans le cône de lumière passé. À la connaissance de l'auteur, c'est le premier exemple d'une solution de ce genre.

Théorème 2.7. *Il existe une solution* $\boldsymbol{u} : (-\infty, T_0] \to \dot{H}^1 \times L^2$ *de* (NLW) *en dimension* $N = 6$ *telle que*

$$\lim_{t \to -\infty} \|\boldsymbol{u}(t) - (\boldsymbol{W} + \boldsymbol{W}_{\frac{1}{\kappa}\mathrm{e}^{-\kappa|t|}})\|_{\dot{H}^1 \times L^2} = 0, \qquad \text{avec } \kappa := \sqrt{\frac{5}{4}}.$$

Il est à noter que $\boldsymbol{u}(t)$ est un exemple d'une solution de (NLW) autre que l'état fondamental qui est globale dans une direction du temps et qui ne contient pas de terme de radiation dans l'espace d'énergie, plus précisément

$$\forall A > 0,\ \limsup_{t \to -\infty} \int_{|x| \geq |t| - A} |\dot{u}(t,x)|^2 + |\nabla u(t,x)|^2 \,\mathrm{d}x = 0.$$

Pour souligner le fait que l'énergie de nos solutions est exactement égale au double de l'énergie de $\boldsymbol{W}$, sans énergie diffusée comme une onde libre, on dit qu'elles sont des *2-bulles pures.*

Donnons quelques commentaires sur la preuve. On observe que dans Théorème 2.1 on peut prendre $\boldsymbol{u}_0^* = (W, 0)$, ce qui produit une solution qui explose en temps fini avec $(W, 0)$ comme profil asymptotique. C'est "presque" ce que l'on désire, sauf que la bulle se concentre et explose en temps fini au lieu d'exister et de se concentrer pour tout temps. Une manière naturelle d'obtenir un tel comportement est d'augmenter la dimension de l'espace. En dimension $N = 6$ l'interaction entre les deux bulles est plus faible, ce qui produit l'effet voulu. Pour $N \geq 7$ on devrait arriver à une conclusion similaire, avec une vitesse de concentration de la bulle $\lambda \sim |t|^{-\frac{4}{N-6}}$.

La principale difficulté technique par rapport au Théorème 2.1 vient du fait que $\boldsymbol{W}$ a une taille fixe dans l'espace d'énergie, alors qu'auparavant, grâce à la vitesse finie de propagation, on a pu supposer que $\|\boldsymbol{u}_0^*\|_{\dot{H}^1 \times L^2}$ était petit. Afin d'obtenir des bornes sur la norme d'énergie, il est maintenant nécessaire d'étudier la coercitivité de la fonctionnelle d'énergie au voisinage d'une somme de deux bulles.

On modifie aussi notre construction pour couvrir le cas de l'équation (WM) avec $k \geq 3$. On obtient le résultat suivant.

Théorème 2.8. *Soit* $k > 2$. *Il existe une solution* $\boldsymbol{u} : (-\infty, T_0] \to \mathcal{E}$ *de* (WM) *telle que*

$$\lim_{t \to -\infty} \|\boldsymbol{u}(t) - (-\boldsymbol{W} + \boldsymbol{W}_{\frac{k-2}{\kappa}(\kappa|t|)^{-\frac{2}{k-2}}})\|_{\mathcal{E}} = 0, \qquad \text{avec } \kappa := \frac{k-2}{2}\Big(\frac{8k}{\pi}\sin\big(\frac{\pi}{k}\big)\Big)^{\frac{1}{k}}.$$

Ici, $\boldsymbol{W}$ désigne l'état fondamental de degré topologique k et $\mathcal{E}$ est l'espace de l'énergie.

Enfin, on démontre un résultat similaire pour l'équation de Yang-Mills critique :

$$\partial_t^2 u(t,r) = \partial_r^2 u(t,r) + \frac{1}{r}\partial_r u(t,r) - \frac{4}{2r^2} u(t,r)(1 - u(t,r))\big(1 - \frac{1}{2}u(t,r)\big).$$

Des résultats de classification montrant la validité de décompositions de type (1.10) ou (1.9) pour une suite de temps $t_n \to T_+$ ont été obtenus par Côte [14] pour l'équation (WM) avec $k = 1$, et par Jia et Kenig [42] en plus grande généralité, y compris dans tous les cas considérés ici.

2.3 Bornes sur la vitesse d'explosion de type II

Les résultats du Chapitre 1 incitent à se pencher davantage sur la relation entre la dynamique de l'explosion et les propriétés du profil asymptotique $\boldsymbol{u}_0^*$. Dans le Chapitre 3, on démontre deux résultats qui vont dans cette direction.

Théorème 2.9. *Soit $N \in \{3,4,5\}$ et $s > \frac{N-2}{2}$, $s \geq 1$. Soit $\boldsymbol{u}_0^* = (u_0^*, \dot{u}_0^*) \in H^{s+1} \times H^s$ à symétrie radiale. Supposons que $\boldsymbol{u}(t)$ est une solution radiale de* (NLW) *telle que*

$$\lim_{t \to T_+} \|\boldsymbol{u}(t) - \boldsymbol{W}_{\lambda(t)} - \boldsymbol{u}_0^*\|_{\dot{H}^1 \times L^2} = 0, \qquad \lim_{t \to T_+} \lambda(t) = 0, \qquad T_+ < +\infty. \tag{2.4}$$

Il existe une constante $C > 0$, qui dépend de $\boldsymbol{u}_0^*$, telle que :

- *si $N \in \{4,5\}$, alors pour $T_+ - t$ suffisamment petit*

$$\lambda(t) \leq C(T_+ - t)^{\frac{4}{6-N}}.$$

- *si $N = 3$, alors il existe une suite $t_n \to T_+$ telle que*

$$\lambda(t_n) \leq C(T_+ - t_n)^{\frac{4}{6-N}}.$$

Théorème 2.10. *Soit $N \in \{3,4,5\}$. Soit $\boldsymbol{u}_0^* = (u_0^*, \dot{u}_0^*) \in H^3 \times H^2$ à symétrie radiale avec*

$$u_0^*(0) < 0.$$

Il n'existe pas de solution $\boldsymbol{u}(t)$ de (NLW) *telle que*

$$\lim_{t \to T_+} \|\boldsymbol{u}(t) - \boldsymbol{W}_{\lambda(t)} - \boldsymbol{u}_0^*\|_{\dot{H}^1 \times L^2} = 0, \qquad \lim_{t \to T_+} \lambda(t) = 0, \qquad T_+ < +\infty. \tag{2.5}$$

L'idée principale des preuves, esquissées dans le Paragraphe 1.4 du Chapitre 3, est de borner le terme d'erreur par l'interaction (énergétique) de la bulle avec le reste de la solution, le principe auquel on a fait allusion dans la Remarque 2.5.

L'image générale qui émerge des Théorèmes 2.1, 2.2 et 2.9 est que la vitesse d'explosion des solutions de type Kriger-Schlag-Tataru est reliée à la régularité du profil asymptotique, plutôt qu'à la régularité de la solution elle-même (même si des solutions qui explosent plus vite ont tendance à être plus régulières, comme indiqué dans [52, 53]). Il faut noter que, même si pour des raisons techniques on demande une régularité supplémentaire de $\boldsymbol{u}_0^*$, dans (2.4) et (2.5) la convergence a lieu seulement dans l'espace d'énergie.

Il est raisonnable de se demander si le profil asymptotique détermine de façon unique la solution qui explose dans un voisinage de $\{\boldsymbol{W}_\lambda\}$. Si c'était le cas, alors en associant à toute solution explosive son profil asymptotique, et réciproquement, on obtiendrait une classification naturelle des solutions explosives au voisinage de $\{\boldsymbol{W}_\lambda\}$ dans l'espace d'énergie. H. Koch, D. Tataru et le rapporteur anonyme de l'article constituant le Chapitre 3 ont indiqué à l'auteur que ce schéma général ressemble au problème classique de la diffusion (scattering). Évidemment on devrait identifier les éléments de l'espace d'énergie qui sont les profils asymptotiques d'une certaine solution explosive. Le contenu du Théorème 2.10 est de montrer des exemples de profils qui doivent être exclus.

Même s'il serait judicieux de regarder le programme décrit ci-dessus avec une certaine réserve, il permet néanmoins de formuler des questions qui semblent plus accessibles. Par exemple, le profil asymptotique détermine-t-il la vitesse de l'explosion ?

2.4 Non existence de 2-bulles de signes opposés

Le Chapitre 4 est consacré à une preuve du résultat suivant :

Théorème 2.11. *Soit* $N \geq 3$. *Il n'existe pas de solution radiale* $\boldsymbol{u} : [t_0, T_+) \to \dot{H}^1 \times L^2$ *de* (NLW) *telle que*

$$\lim_{t \to T_+} \|\boldsymbol{u}(t) - \boldsymbol{W}_{\lambda(t)} + \boldsymbol{W}_{\mu(t)}\|_{\dot{H}^1 \times L^2} = 0$$

et

- *dans le cas* $T_+ < +\infty$, $\lambda(t) \ll \mu(t) \ll T_+ - t$ *quand* $t \to T_+$,
- *dans le cas* $T_+ = +\infty$, $\lambda(t) \ll \mu(t) \ll t$ *quand* $t \to +\infty$.

Comme dans le contexte du Théorème 2.7, on travaille ici au niveau énergétique $2E(W)$, qui est le seuil d'énergie pour une formation d'une multi-bulle. La motivation principale provient de l'équation (WM), où les 2-bulles peuvent soit avoir les orientations contraires, auquel cas le degré topologique de la solution vaut 0, soit avoir la même orientation, ce qui impliquerait que le degré topologique de la solution vaut $\pm 2k$. La première situation est l'objet du Théorème 2.8, et l'on sait qu'une deux-bulle peut se former, au moins dans le cas $k \geq 3$. D'un autre côté, par des arguments variationnels bien connus, l'énergie potentielle de toute application k-équivariante de degré topologique $2k$ excède $2E(W)$, donc la conservation de l'énergie implique qu'il n'y a pas de 2-bulles pures dans ce dernier cas.

Notre preuve du Théorème 2.11 est de nature variationnelle, tout comme la preuve dans le cas des applications harmoniques décrite ci-dessus. Le contenu dynamique y est réduit à l'étude de la dynamique hyperbolique induite par la présence d'une direction stable et d'une direction instable du flot au voisinage de $\boldsymbol{W}$, ce qui constitue une différence majeure par rapport à l'équation (WM). La partie la plus difficile est d'exclure l'existence de solutions globales qui se comporteraient asymptotiquement comme une superposition d'une bulle positive à l'échelle 1 et d'une bulle négative qui se concentre. À cette fin, en utilisant *la variété stable* construite par Duyckaerts et Merle [27], nous définissons des directions d'instabilité raffinées (non linéaires), qui permettent d'obtenir des propriétés de coercitivité de la fonctionnelle d'énergie suffisamment précises et robustes. Ce schéma de preuve est influencé par les résultats de Krieger, Nakanishi et Schlag [49], que l'on a déjà mentionnés dans le Paragraphe 1.4.

Chapter II

Introduction (english version)

1 Generalities on nonlinear wave equations

The subject of this thesis is to study some nonlinear phenomena in the dynamical behavior of radial solutions of the focusing wave equation with the energy-critical power nonlinearity in dimension $N \geq 3$:

$$\partial_t^2 u(t,x) = \Delta u(t,x) + |u(t,x)|^{\frac{4}{N-2}} u(t,x), \qquad (t,x) \in \mathbb{R} \times \mathbb{R}^N, \tag{NLW}$$

and of k-equivariant solutions of the wave map equation from $\mathbb{R}^{1+2}$ to the two-dimensional sphere $\mathbb{S}^2$:

$$\partial_t^2 u(t,r) = \partial_r^2 u(t,r) + \frac{1}{r}\partial_r u(t,r) - \frac{k^2}{2r^2}\sin(2u(t,r)), \qquad (t,r) \in \mathbb{R} \times (0,+\infty). \tag{WM}$$

Before stating the main results, let us briefly present the subject of the global theory of nonlinear wave equations from a historical perspective.

1.1 Preliminaries

The simplest nonlinear wave equations are of the form

$$\begin{cases} \partial_t^2 u(t,x) = \Delta u(t,x) + f(x, u(t,x)), & (t,x) \in \mathbb{R} \times \mathbb{R}^N, \\ u(t_0,x) = u_0(x),\ \partial_t u(t_0,x) = \dot u_0(x), & x \in \mathbb{R}^N, \end{cases} \tag{1.1}$$

where Δ is the Laplacian in the N space variables and f is a (smooth enough) function of 2 real variables such that $f(x,0) = 0$. This equation has a natural Hamiltonian structure. Indeed, let $F(x,u) := \int_0^u f(x,v)\,\mathrm{d}v$ and define the *energy functional*

$$\begin{aligned} &E : C_0^\infty(\mathbb{R}^N) \times C_0^\infty(\mathbb{R}^N) \ni \boldsymbol{u} = (u, \dot u) \mapsto E(\boldsymbol{u}) \in \mathbb{R}, \\ &E(\boldsymbol{u}) := \int_{\mathbb{R}^N} \frac{1}{2}|\dot u(x)|^2 + \frac{1}{2}|\nabla u(x)|^2 - F(x, u(x))\,\mathrm{d}x. \end{aligned}$$

Equation (1.1) can be written in the form

$$\begin{cases} \partial_t \boldsymbol{u} = J \circ \mathrm{D}E(\boldsymbol{u}), \\ \boldsymbol{u}(t_0) = \boldsymbol{u}_0, \end{cases} \tag{1.2}$$

where D is the Fréchet derivative, $J := \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix}$ defines the natural symplectic form on $C_0^\infty \times C_0^\infty$ and $\boldsymbol{u} = (u, \dot u)$ is an element of the phase space $C_0^\infty \times C_0^\infty$.

Denote $\langle \boldsymbol{u}, \boldsymbol{v}\rangle := \int_{\mathbb{R}^N} u\cdot v + \dot u \cdot \dot v\,\mathrm{d}x$. Let $\boldsymbol{u}(t)$ be a classical solution of (1.2) on a time interval (t_1, t_2). The chain rule yields

$$\frac{\mathrm{d}}{\mathrm{d}t}E(\boldsymbol{u}(t)) = \langle \mathrm{D}E(\boldsymbol{u}(t)), \partial_t \boldsymbol{u}(t)\rangle = \langle \mathrm{D}E(\boldsymbol{u}(t)), J\circ \mathrm{D}E(\boldsymbol{u}(t))\rangle = 0, \tag{1.3}$$

hence $E(\boldsymbol{u}(t))$ is a conserved quantity.

1.2 A few historical remarks

The history of the local Cauchy problem for the wave equations goes back at least to the Cauchy-Kowalevski theorem, which, for analytic nonlinearities, yields local existence and uniqueness of an analytic solution for analytic initial data. However, from the physical point of view the question of global existence is highly relevant, and the first result in this direction is due to Jörgens [43], followed by more abstract works of Browder [7] and Segal [83]. Let us state the result in the simplest case of the cubic defocusing Klein-Gordon equation in $\mathbb{R}\times\mathbb{R}^3$, that is equation (1.1) with $f(x,u) = -u-u^3$:

$$\partial_t^2 u(t,x) = \Delta u(t,x) - u(t,x) - u(t,x)^3, \qquad (t,x)\in\mathbb{R}\times\mathbb{R}^3. \tag{1.4}$$

Theorem 1.1 (Jörgens, Browder, Segal)**.** *For any initial data $\boldsymbol{u}(t_0) = \boldsymbol{u}_0 \in H^1(\mathbb{R}^3)\times L^2(\mathbb{R}^3)$, there exists a unique global solution $\boldsymbol{u}(t)$ of* (1.4)*. Moreover, if $\boldsymbol{u}_0\in H^{k+1}\times H^k$ with $k\in\mathbb{N}$, then $\boldsymbol{u}\in C(\mathbb{R}; H^{k+1}\times H^k)$.*

Remark 1.2. Later in this section we will clarify the notion of a solution in the non-smooth setting.

The main ingredient of the proof is the following local existence result.

Proposition 1.3. *For any initial data $\boldsymbol{u}(t_0) = \boldsymbol{u}_0\in H^1(\mathbb{R}^3)\times L^2(\mathbb{R}^3)$, there exists a unique maximal solution $\boldsymbol{u}(t) : (T_-, T_+)\to H^1\times L^2$ of* (1.4)*. Moreover, if $\boldsymbol{u}_0\in H^{k+1}\times H^k$ with $k\in\mathbb{N}$, then $\boldsymbol{u}\in C((T_-,T_+); H^{k+1}\times H^k)$. If $T_+ < +\infty$, then $\lim_{t\to T_+}\|\boldsymbol{u}(t)\|_{H^1\times L^2} = +\infty$ (analogously if $T_- > -\infty$). Finally, on compact time intervals the solution is continuous with respect to the initial data in the topology $H^{k+1}\times H^k$ for all $k\in\mathbb{N}$.*

By (1.3) and a standard approximation procedure one obtains that if $\boldsymbol{u}(t)$ is a solution given by Proposition 1.3, then the energy

$$E(\boldsymbol{u}(t)) = \int_{\mathbb{R}^N}\frac12|\dot u(t,x)|^2 + \frac12|\nabla u(t,x)|^2 + \frac12|u(t,x)|^2 + \frac14|u(t,x)|^4\,\mathrm{d}x$$

is constant. But this implies that $\|\boldsymbol{u}(t)\|_{H^1\times L^2}$ is bounded, thus proving Theorem 1.1.

Since the nonlinearity is defocusing, it was conjectured that the solution has to decay as $t\to\pm\infty$ at least like the free solutions, causing the nonlinear effects to become negligible. Thus, any solution should behave asymptotically like a solution of a free Klein-Gordon equation (it *scatters*), in other words one expects asymptotic completeness in an appropriate space. This problem turned out to be quite challenging and was solved in the '80 by Brenner [6], with essential earlier contributions of Morawetz and Strauss [70].

Theorem 1.4 (Brenner, Morawetz, Strauss)**.** *Let $\boldsymbol{u}(t)$ be a solution of* (1.4) *and let $U_0(t)$ be the linear Klein-Gordon propagator. Then $U_0(-t)\boldsymbol{u}(t)$ has a strong limit in $H^1\times L^2$ as $t\to\pm\infty$.*

Remark 1.5. Existence of a weak limit is relatively elementary and was already known in the '60.

Oversimplifying, one could say that the dynamical behavior of the solutions of (1.4) does not differ much from the free dynamics, at least in the energy space. In establishing these results, the sign of the nonlinearity (repulsive) and its growth at $\pm\infty$ (u^3) are decisive. It is natural to examine the case of a focusing nonlinearity or of a faster growth of the nonlinearity (for example u^5), which will be briefly discussed in the next two paragraphs.

1.3 Effects of a focusing nonlinearity

Consider equation (1.1) with $f(x, u) = -u + u^3$ (the *focusing* cubic Klein-Gordon equation), in dimension $N = 3$:

$$\partial_t^2 u(t, x) = \Delta u(t, x) - u(t, x) + u(t, x)^3, \qquad (t, x) \in \mathbb{R} \times \mathbb{R}^3. \tag{1.5}$$

Nothing changes as far as the local Cauchy theory is concerned (Proposition 1.3 remains valid). However, the energy functional corresponding to the focusing case is

$$E(\boldsymbol{u}) = \int_{\mathbb{R}^3} \frac{1}{2}|\dot{u}|^2 + \frac{1}{2}|\nabla u|^2 + \frac{1}{2}|u|^2 - \frac{1}{4}|u|^4 \,\mathrm{d}x,$$

hence it no longer allows to control the $H^1 \times L^2$ norm for large times. The fact that Theorem 1.1 fails in the focusing case was first observed by Keller [45]. We see that sufficiently large constant in space initial data lead to solutions tending to $+\infty$ in finite time. Using the finite speed of propagation, this produces initial data in C_0^∞ which lead to a blow-up in finite time.

Another consequence of the focusing sign is the existence of stationary solutions W, that is solutions of the elliptic problem

$$\Delta W(x) - W(x) + W(x)^3 = 0, \qquad x \in \mathbb{R}^3. \tag{1.6}$$

The initial data $\boldsymbol{u}_0 := \boldsymbol{W} := (W, 0)$ lead to a constant in time solution of (1.5). Among the stationary solutions, the *ground state* plays a distinguished role. It can be characterized as the unique (up to translations) positive solution of (1.6), or as the solution minimizing the potential energy $E(u) := \int_{\mathbb{R}^3} \frac{1}{2}|\nabla u|^2 + \frac{1}{2}|u|^2 - \frac{1}{4}|u|^4 \,\mathrm{d}x$. The variational picture is that for u small in the energy norm $E(u)$ creates a potential well and W acts as the mountain pass. Payne and Sattinger [75] exhibited the role of $E(W)$ as the *threshold energy*. Namely, if a solution of energy $E(\boldsymbol{u}_0) < E(W)$ is inside the potential well ($\|u_0\|_{H^1} \leq \|W\|_{H^1}$), it stays there forever, and if it is outside ($\|u_0\|_{H^1} > \|W\|_{H^1}$), it blows up in finite time.

Since the solutions in the potential well are global, it is natural to study if they converge to linear solutions in the energy space, as in the defocusing case. A general approach to this problem, called the *Ground State Conjecture* or the *Threshold Conjecture*, was developed by Kenig and Merle [46, 47] for the energy-critical wave equation and the energy-critial radial Schrödinger equation. In the case of the cubic Klein-Gordon equation, the proof was carried out by Ibrahim, Masmoudi and Nakanishi [38].

Theorem 1.6 (Ibrahim, Masmoudi, Nakanishi)**.** *Let $\boldsymbol{u}(t)$ be a solution of* (1.5) *such that $E(\boldsymbol{u}_0) < E(W)$ and $\|\boldsymbol{u}_0\|_{H^1 \times L^2} \leq \|W\|_{H^1}$. Then $U_0(-t)\boldsymbol{u}(t)$ has a strong limit in $H^1 \times L^2$ as $t \to \pm\infty$.*

The next crucial issue is to understand the dynamical behavior *slightly above* the ground state energy. A very elegant solution of this problem, showing a link with the theory of invariant manifolds, was given by Nakanishi and Schlag [71, 72].

1.4 Energy-critical nonlinearities

Let us return for a moment to the general equation (1.1). We have seen that the structure of the problem often allows to control directly or indirectly the homogeneous energy norm

$$\|\boldsymbol{u}(t)\|_{\dot H^1\times L^2}^2 := \int_{\mathbb{R}^N} |\partial_t u(t,x)|^2 + |\nabla u(t,x)|^2\,\mathrm{d}x,$$

giving some information on the regularity of a solution (in the Klein-Gordon case we can control also $\|u\|_{L^2}$, but this is less interesting from the point of view of regularity). Whether this is enough to preclude a breakdown of a solution will depend on the nonlinearity, and more specifically on the growth of $|f(x,u)|$ as $|u|\to+\infty$. We will study heuristically the possibility of a singularity formation such that the energy norm remains bounded. This type of singularity is referred to as a *type II blow-up* or a *geometric blow-up* (in opposition to a *type I blow-up*, referring to a situation where the energy norm tends to $+\infty$ in finite time).

For $\boldsymbol{u}(x) = (u(x),\dot u(x))$ and $\lambda>0$ we define the energy-critical scale change

$$u_\lambda(x) := \frac{1}{\lambda^{\frac{N-2}{2}}}u\Big(\frac{x}{\lambda}\Big),\quad \boldsymbol{u}_\lambda(x) := \Big(\frac{1}{\lambda^{\frac{N-2}{2}}}u\Big(\frac{x}{\lambda}\Big),\frac{1}{\lambda^{\frac N2}}\dot u\Big(\frac{x}{\lambda}\Big)\Big)$$

A straightforward change of variables shows that $\|\boldsymbol{u}_\lambda\|_{\dot H^1\times L^2} = \|\boldsymbol{u}\|_{\dot H^1\times L^2}$. At the same time, the energy norm *concentrates at the origin* as $\lambda\to 0$. Suppose to simplify that $f(x,u)=f(u)$ does not depend on x and that $N\geq 3$. If we assume that $|f(u)|\lesssim |u|^p$ as $|u|\to+\infty$, then

$$|f(u_\lambda)|\lesssim \frac{1}{\lambda^{\frac{p(N-2)}{2}-\frac{N-2}{2}}}\big|\big(|u|^p\big)_\lambda\big|,\quad \text{for small } \lambda,$$

whereas

$$\Delta u_\lambda = \frac{1}{\lambda^2}\big(\Delta u\big)_\lambda.$$

If $\frac{p(N-2)}{2}-\frac{N-2}{2}<2\Leftrightarrow p<\frac{N+2}{N-2}$, then the effect of the nonlinearity is negligible as compared to the laplacian as $\lambda\to 0$. Intuitively, concentration of the energy norm causes the nonlinear effects to become negligible, in particular cannot result in a breakdown of the solution. The only way of forming a singularity is thus the growth of the energy norm, which is a part of Proposition 1.3. This type of nonlinearity is called *energy-subcritical.* If, on the contrary, $p>\frac{N+2}{N-2}$, then the nonlinearity is the dominant part as $\lambda\to 0$. This means that a concentration of an arbitrarily small amount of energy could potentially cause the nonlinear effects to prevail, resulting in a singularity. Such a nonlinearity is called *energy-supercritical.* The borderline case, $p=\frac{N+2}{N-2}$, is called *energy-critical.* In this last situation, as the data concentrates, the nonlinearity acts with the same force as the laplacian, a priori leading to a finite-energy singularity.

Probably the simplest energy-critical wave equation is the semilinear defocusing wave equation with the energy-critical power nonlinearity:

$$\partial_t^2 u(t,x) = \Delta u(t,x) - |u(t,x)|^{\frac{4}{N-2}}u(t,x),\qquad (t,x)\in\mathbb{R}\times\mathbb{R}^N. \tag{1.7}$$

The local well-posedness in the energy space $\dot H\times L^2$ is due to Shatah and Struwe [84], with important contributions of Kapitanski [44] and Ginibre, Soffer and Velo [32]. Note that for $\boldsymbol{u}=(u,\dot u)\in\dot H^1\times L^2$ the energy $E(\boldsymbol{u}) = \int_{\mathbb{R}^N}\frac12|\dot u|^2+\frac12|\nabla u|^2+\frac{N-2}{2N}|u|^{\frac{2N}{N-2}}\,\mathrm{d}x$ is well defined due to the critical Sobolev embedding. We also have $E(\boldsymbol{u}_\lambda)=E(\boldsymbol{u})$, because the nonlinearity has the same scaling as the linear part. Finally, the equation is invariant by the energy-critical change of scale: if $(u(t,x),\dot u(t,x))$ is a solution of (1.7) on an interval $I\ni t_0$, then for any

$\lambda > 0$ we can build another solution $\big(u\big(t_0 + \frac{t}{\lambda}, \frac{x}{\lambda}\big)\big)$, defined on some time interval containing $t = 0$.

It turns out that, for the solutions of (1.7), the energy norm cannot concentrate, and that an analog of Theorem 1.1 holds, which was proved by Struwe [88] in the radial case and by Grillakis [34] without the symmetry assumption. Also, the solutions scatter, i.e. asymptotically approach linear waves in the energy space as $t \to \pm\infty$. In [91, Chapter 5] one can find some details about the history of the resolution of this problem.

Most of this thesis is concerned with equation (NLW), which is the focusing counterpart of (1.7). The energy is given by

$$E(\boldsymbol{u}) = \int_{\mathbb{R}^N} \frac{1}{2}|\dot u|^2 + \frac{1}{2}|\nabla u|^2 - \frac{N-2}{2N}|u|^{\frac{2N}{N-2}}\,\mathrm{d}x.$$

The local well-posedness can be established by modifying the proofs in the defocusing case, see [46]. One obtains the following result.

Proposition 1.7. *For any initial data $\boldsymbol{u}_0 \in \dot H^1(\mathbb{R}^N)\times L^2(\mathbb{R}^N)$, there exists a unique maximal solution $\boldsymbol{u}(t) : (T_-, T_+) \to \dot H^1 \times L^2$ of* (NLW)*. If $\|\boldsymbol{u}_0\|_{\dot H^1\times L^2}$ is sufficiently small, then $T_- = -\infty$, $T_+ = +\infty$ and $\boldsymbol{u}(t)$ scatters in both time directions. The solution is continuous with respect to the initial data in the topology $\dot H^1 \times L^2$.*

Note an important difference with respect to the sub-critical case discussed above. If $T_+ < +\infty$, it is no longer guaranteed that $\lim_{t\to T_+} \|\boldsymbol{u}\|_{\dot H^1\times L^2} = +\infty$. The reason for this was heuristically described at the beginning of this paragraph. We have however the standard abstract blow-up criterion: if $T_+ < +\infty$ and K is a compact subset of $\dot H^1 \times L^2$, then there exists $\tau > 0$ such that $\boldsymbol{u}(t) \notin K$ for $t \in [T_+ - \tau, T_+)$.

The elliptic problem

$$\Delta W(x) + |W(x)|^{\frac{4}{N-2}}W(x) = 0, \qquad x \in \mathbb{R}^N \tag{1.8}$$

(called the Yamabe equation) has an explicit positive solution

$$W(x) = \Big(1 + \frac{|x|^2}{N(N-2)}\Big)^{-\frac{N-2}{2}}.$$

Because of the energy-critical scaling invariance, for all $\lambda > 0$, $W_\lambda(x)$ is also a solution of (1.8). It is known that these are the only finite-energy radially symmetric and, up to translations, the only positive solutions of (1.8). However, the nonradial problem is complicated, see [73]. We obtain the radially symmetric stationary solutions of (NLW) $\boldsymbol{W}_\lambda = (W_\lambda, 0)$. The role of W_λ as the mountain passes for the potential energy $E(u) = \int \frac{1}{2}|\nabla u|^2 - \frac{N-2}{2N}|u|^{\frac{2N}{N-2}}\,\mathrm{d}x$ follows from the works of Aubin [1] and Talenti [90].

The Ground State Conjecture for the energy-critical focusing wave equation was proved in the work of Kenig and Merle [46] for $N \in \{3, 4, 5\}$. For $N = 3$, Krieger, Schlag and Tataru [53] constructed solutions developing a singularity in finite time, while staying in an arbitrarily small neighborhood (in the energy space) of the family of the ground states $\boldsymbol{W}_\lambda$. In particular, this construction gave the first example of a type II blow-up. Duyckaerts, Kenig and Merle [23, 25] proved that *any* solution of (NLW) blowing up in a neighborhood of the set $\{\boldsymbol{W}_\lambda\}$ asymptotically decomposes in the energy space as a sum of a ground state, rescaled and Lorentz-transformed, and an asymptotic profile. More precisely, in the case of radial data, if $T_+ < +\infty$, then there exists a positive function $\lambda(t) \ll T_+ - t$ and $\boldsymbol{u}_0^* \in \dot H^1 \times L^2$ such that

$$\lim_{t\to T_+} \|\boldsymbol{u}(t) - \boldsymbol{u}_0^* - \boldsymbol{W}_{\lambda(t)}\|_{\dot H^1\times L^2} = 0.$$

In dimension $N = 3$ for radially symmetric data, the same authors [26] gave a complete classification of the dynamical behavior of solutions of (NLW) in the energy space. They proved that if $T_+ = +\infty$, then there exists $J \in \mathbb{N}$, positive functions $\lambda_1(t) \ll \lambda_2(t) \ll \ldots \lambda_J(t) \ll t$ and a solution $\boldsymbol{u}^*(t)$ of the linear wave equation such that

$$\lim_{t\to+\infty} \Big\|\boldsymbol{u}(t) - \boldsymbol{u}^*(t) - \sum_{j=1}^{J} \pm \boldsymbol{W}_{\lambda_j(t)}\Big\|_{\dot H^1\times L^2} = 0. \tag{1.9}$$

If $T_+ < +\infty$, then either $\lim_{t\to T_+} \|\boldsymbol{u}(t)\|_{\dot H^1\times L^2} = +\infty$, or there exists $J \in \mathbb{N}$, positive functions $\lambda_1(t) \ll \lambda_2(t) \ll \ldots \lambda_J(t) \ll T_+ - t$ and $\boldsymbol{u}_0^* \in \dot H \times L^2$ such that

$$\lim_{t\to T_+} \Big\|\boldsymbol{u}(t) - \boldsymbol{u}_0^* - \sum_{j=1}^{J} \pm \boldsymbol{W}_{\lambda_j(t)}\Big\|_{\dot H^1\times L^2} = 0. \tag{1.10}$$

Krieger, Nakanishi and Schlag [49] studied the dynamics in a neighborhood of the ground states from a different perspective, establishing a connection with the theory of invariant manifolds. In particular, they construct a C^1 hypersurface containing $\{\boldsymbol{W}_\lambda\}$, the *center-stable manifold*, and show that it separates the set of solutions which scatter forward in time from the solutions which develop a type I singularity. The solutions on the manifold are the ones which stay close to $\{\boldsymbol{W}_\lambda\}$ as long as they exist.

Remark 1.8. A global theory of energy-supercritical wave equations for large data, even in the defocusing case, currently seems to be out of reach. To the author's knowledge, only conditional results about the dynamics of solutions are available (both in the focusing and defocusing case), see Kenig and Merle [48]. In the case of a defocusing power nonlinearity, Krieger and Schlag [51] proved existence of sets of large initial data leading to global, regular solutions.

1.5 Other critical models

Historically, the most intensively studied energy-critical model is the *harmonic map heat flow* between surfaces, especially between two-dimensional spheres. For a state of the art in this domain, see Topping [94] and the references therein. In particular, the first results classifying the blow-up behavior (*bubbling*) were obtained in this case.

The terminology "type I blow-up" and "type II blow-up" originates in the works on energy-supercritical heat equation $\partial_t u = \Delta u + |u|^p u$, see [63].

An intensively studied model, relevant from the physical point of view, is the mass-critical (or L^2-critical) focusing Schrödinger equation. In some aspects it resembles the L^2-critical generalized Korteweg-de Vries equation. The expository article [61] presents the recent major advances in the understanding of these two equations.

Important energy-critical Hamiltonian equations include the nonlinear Schrödinger equation: $i\partial_t u + \Delta u + |u|^{\frac{4}{N-2}}u = 0$ on $\mathbb{R}^N$, and the Schrödinger map equation from $\mathbb{R}^2$ to the sphere $\mathbb{S}^2 \subset \mathbb{R}^3$: $\partial_t u = u \wedge \Delta u$, where $\wedge$ is the vector product in $\mathbb{R}^3$. In the sequel we will occasionally mention these models, for the sake of comparison with the nonlinear wave equations.

Remark 1.9. Energy critical equations usually generate complicated dynamical behavior, but they also come with a special additional structure of the invariance with respect to the energy-critical rescaling, which is crucial in many arguments.

1.6 Rudiments of the Cauchy theory

For the sake of completeness, in this paragraph we comment on the notion of a solution and of well-posedness. It will be convenient to denote $X^s := \dot{H}^{s+1} \cap \dot{H}^1$ pour $s \geq 0$.

Definition 1.10. Let $t_0 \in (t_1, t_2) \subset \mathbb{R}$, $\boldsymbol{u}_0 = (u_0, \dot{u}_0) \in X^s \times H^s$ and $h \in L^1((t_1, t_2); H^s)$. For $t \in (t_1, t_2)$ set

$$\begin{aligned} u(t) &:= \cos\big((t - t_0)\sqrt{-\Delta}\big)u_0 + \frac{\sin\big((t - t_0)\sqrt{-\Delta}\big)}{\sqrt{-\Delta}}\dot{u}_0 \\ &\quad + \int_{t_0}^t \frac{\sin\big((t - s)\sqrt{-\Delta}\big)}{\sqrt{-\Delta}} h(s)\,\mathrm{d}s, \\ \dot{u}(t) &:= -\sin\big((t - t_0)\sqrt{-\Delta}\big)\sqrt{-\Delta}u_0 + \cos\big((t - t_0)\sqrt{-\Delta}\big)\dot{u}_0 \\ &\quad + \int_{t_0}^t \cos\big((t - s)\sqrt{-\Delta}\big)h(s)\,\mathrm{d}s. \end{aligned}$$

We call $\boldsymbol{u}(t) := (u(t), \dot{u}(t))$ the *solution* of the problem

$$\begin{cases} \partial_t^2 u = \Delta u + h, \\ (u, \partial_t u)_{t=t_0} = (u_0, \dot{u}_0). \end{cases} \tag{1.11}$$

Note that it follows directly from the formulas that $\boldsymbol{u} \in C((t_1, t_2), X^s \times H^s)$ (which is commonly expressed by saying that "the wave equation gains one derivative"). A quantitative version of this fact is the following fundamental *energy estimate*:

$$\|\boldsymbol{u}(t)\|_{X^s \times H^s} \leq \|\boldsymbol{u}_0\|_{X^s \times H^s} + \Big|\int_{t_0}^t \|h(s)\|_{H^s}\,\mathrm{d}s\Big|. \tag{1.12}$$

Existence and uniqueness of weak solutions of the linear problem (1.11), as well as the energy estimate (in the setting of more general symmetric systems) are due to Friedrichs [30].

Definition 1.11. Suppose that $\partial_u f(x, 0) = 0$. We say that (1.1) is *locally well-posed* in $X^s \times H^s$ if

- for any $\boldsymbol{u}_0 \in X^s \times H^s$ and $t_0 \in \mathbb{R}$ there exists a time interval $(t_1, t_2) \ni t_0$ and a unique $\boldsymbol{u}(t) = (u(t), \dot{u}(t))$ such that

$$f(x, u(t, x)) \in L^1((t_1, t_2); H^s) \tag{1.13}$$

  and $\boldsymbol{u}(t)$ is the solution of (1.11) with $h(t, x) = f(x, u(t, x))$,
- $\boldsymbol{u}(t)$ depends continuously in $X^s \times H^s$ on $\boldsymbol{u}_0$.

Remark 1.12. For smooth functions, this definition agrees with the usual definition of a classical solution of a differential equation.

Remark 1.13. In general one should consider the operator $T := -\Delta - \partial_u f(x, 0)$ instead of $-\Delta$. Under some natural assumptions on T, we can define a solution in the same manner, replacing everywhere H^s by $(1 + T^{\frac{s}{2}})^{-1}L^2$ and X^s by $(T^{\frac{1}{2}} + T^{\frac{s+1}{2}})^{-1}L^2$.

Remark 1.14. In typical situations the condition (1.13) is equivalent to $u(t)$ being in some natural function space, for example a Lebesgue space on a slab of the space-time. Definition 1.11 describes a fixed-point problem. The main part of its resolution is to obtain bounds of $\|f(x, u(t, x))\|_{L^1((t_1,t_2);H^s)}$ in terms of $\|h(t, x)\|_{L^1((t_1,t_2);H^s)}$, where $u(t, x)$ is the solution of (1.11). To this end, one uses Sobolev, Strichartz and other inequalities of this type.

Remark 1.15. Suppose that (1.1) is locally well-posed in $X^s \times H^s$ and that $\boldsymbol{u}(t)$ is a solution with maximal time of existence $T_+ < +\infty$. Then $\|f(x, u(t,x))\|_{L^1([t_0,T_+);H^s)} = +\infty$. Otherwise, as follows from (1.12), the solution $\boldsymbol{u}(t)$ would be precompact in $X^s \times H^s$ as $t \to T_+$, which is impossible.

In the same vain, if $T_+ = +\infty$ and $\|f(x, u(t,x))\|_{L^1([t_0,+\infty);H^s)} < +\infty$, then $\boldsymbol{u}(t)$ scatters in $X^s \times H^s$ as $t \to +\infty$.

Remark 1.16. In some more "pathological" cases, such as (NLW) for large N, this general definition might not be the correct one, and it should be settled in each individual case what additional condition on $\boldsymbol{u}$, instead of $f(x, u(t,x)) \in L^1((t_1, t_2); H^s)$, ensures that the solution is unique and continuously dependent on the initial data.

2 Main results

In this thesis we consider equation (NLW) for radially symmetric initial data, except in Chapter 2, where we deal also with equation (WM).

Following the works mentioned above, we undertake the study of the local behavior of the dynamical system defined by (NLW) in a small neighborhood of $\{\boldsymbol{W}_\lambda\}$ in the energy space, that is the solutions $\boldsymbol{u}(t)$ of (NLW) such that

$$\inf_{\lambda>0} \|\boldsymbol{u}(t) - \boldsymbol{W}_\lambda\|_{\dot{H}^1\times L^2} \leq \eta, \qquad \forall t,$$

where $\eta > 0$ is a small constant. Chapters 1 and 3 are devoted to the phenomenon of the type II blow-up.

In Chapters 2 and 4 we study the local behavior in a neighborhood of a superposition of two bubbles at different scales, that is solutions $\boldsymbol{u}(t)$ of (NLW) such that

$$\inf_{0<\lambda\leq\alpha^*\mu} \|\boldsymbol{u}(t) - (\boldsymbol{W}_\mu \pm \boldsymbol{W}_\lambda)\|_{\dot{H}^1\times L^2} \leq \eta, \qquad \forall t,$$

where $\eta > 0$ is a small constant. The main motivation for this work are of course the classification results of [26] mentioned above, see (1.9) and (1.10).

This thesis is composed of the following four articles:

- Chapter 1 – Jendrej, J. Construction of type II blow-up solutions for the energy-critical wave equation in dimension 5. *Preprint*, arXiv:1503.05024, 2015.
- Chapter 2 – Jendrej, J. Construction of two-bubble solutions for energy-critical wave equations. *Preprint*, arXiv:1602.06524, 2016.
- Chapter 3 – Jendrej, J. Bounds on the speed of type II blow-up for the energy critical wave equation in the radial case. *Int. Math. Res. Not.*, doi: 10.1093/imrn/rnv365, 2015.
- Chapter 4 – Jendrej, J. Nonexistence of radial two-bubbles with opposite signs for the energy-critical wave equation. *Preprint*, arXiv:1510.03965, 2015.

2.1 Construction of type II blow-up solutions

In this article we propose a new approach to constructing type II blow-up solutions for the energy-critical wave equation. Such solutions were first constructed for the energy-critical wave equation in dimension $N = 3$ and for the energy-critical wave map equation in equivariance class $k = 1$ by Krieger, Schlag and Tataru [52, 53].

Here, we consider (NLW) in dimension $N = 5$. We prove the following results.

Theorem 2.1. *Let* $\boldsymbol{u}_0^* = (u_0^*, \dot u_0^*) \in (\dot H^5 \cap \dot H^1) \times H^4$ *be any pair of radial functions such that* $u_0^*(0) > 0$. *There exists a solution* $\boldsymbol{u}(t)$ *of* (NLW) *defined on a time interval* $(0, T_0)$ *such that*

$$\lim_{t\to 0} \big\|\boldsymbol{u}(t) - \boldsymbol{u}_0^* - \boldsymbol{W}_{\lambda(t)}\big\|_{\dot H^1\times L^2} = 0, \tag{2.1}$$

where $\lambda(t) = \big(\frac{32}{315\pi}\big)^2\big(u^*(0,0)\big)^2 t^4$.

Theorem 2.2. *Let* $\nu > 8$. *There exists a solution* $\boldsymbol{u}(t)$ *of* (NLW) *defined on the time interval* $(0, T_0)$ *such that*

$$\lim_{t\to 0} \big\|\boldsymbol{u}(t) - \boldsymbol{u}_0^* - \boldsymbol{W}_{\lambda(t)}\big\|_{\dot H^1\times L^2} = 0,$$

where $\lambda(t) = t^{\nu+1}$ *and* $\boldsymbol{u}_0^*$ *is explicit.*

The fact that we obtain a continuous range of blow-up speeds leads us to believe that the dynamical behavior of our solutions is highly unstable with respect to variations of the initial data. A bit more precisely, we can expect that the set of solutions having the same dynamical characteristics (for example the same blow-up rate) has infinite codimension. We could say that we obtain solutions of Krieger-Schlag-Tataru type, in contrast with the *stable* solutions constructed by Hillairet and Raphaël [36] in dimension $N = 4$, which can blow up only at specific rates. Such solutions were also constructed for (WM) by Rodnianski and Sterbenz [80], and Raphaël and Rodnianski [79].

Remark 2.3. In this second case numerical experiments suggested the possibility of a formation of singularities, see for instance [4]. The blow-up observed numerically is the stable one. The authors of [4] suggest that the set of initial data leading to a singularity is huge. However, existence of an open subset of the energy space such that any initial data in this subset develops a singularity, remains unknown.

Let us present the main ideas of the proof of Theorem 2.1, ignoring all the technical issues.

Step 1. Suppose that $\boldsymbol{u}(t) = (u(t), \dot u(t))$ satisfies (2.1). Let $\boldsymbol{u}^*(t) = (u^*(t), \dot u^*(t))$ be the solution of (NLW) with the initial data $\boldsymbol{u}^*(0) = \boldsymbol{u}_0^*$. Denote $\Lambda W := -\frac{\mathrm{d}}{\mathrm{d}\lambda} W_\lambda|_{\lambda=1}$. The chain rule yields $\partial_t W_{\lambda(t)} = -\frac{\lambda'(t)}{\lambda(t)}\Lambda W_{\lambda(t)}$, hence we should have

$$(u(t), \dot u(t)) \simeq \Big(u^*(t) + W_{\lambda(t)}, \dot u^*(t) - \frac{\lambda'(t)}{\lambda(t)}\Lambda W_{\lambda(t)}\Big). \tag{2.2}$$

We find that it is a natural choice to work in the energy space. By a change of variable $\big\|\frac{\lambda'(t)}{\lambda(t)}\Lambda W_{\lambda(t)}\big\|_{L^2} \sim \lambda'(t)$ (the solution explodes at $t = 0$, thus $\lambda'(t) > 0$). If we introduce a small parameter $b(t) \simeq \lambda'(t)$, (2.2) can be seen as a beginning of an asymptotic expansion of $\boldsymbol{u}(t)$ in the energy space in powers of $b(t)$.

It turns out that if

$$b'(t) = \frac{128}{105\pi} u^*(t)\sqrt{\lambda(t)}, \tag{2.3}$$

then we can compute the next term of the expansion and define an *approximate solution* (which we will also call an *ansatz*)

$$\boldsymbol{\varphi}(t) := \boldsymbol{u}^*(t) + \boldsymbol{U}^{(0)}_{\lambda(t)} + b(t)\boldsymbol{U}^{(1)}_{\lambda(t)} + b^2(t)\boldsymbol{U}^{(2)}_{\lambda(t)} + b^3(t)\boldsymbol{U}^{(3)}_{\lambda(t)},$$

where $\boldsymbol{U}^{(0)} = (W, 0)$ and $\boldsymbol{U}^{(1)} = (0, -\Lambda W)$. The precise definitions of $\boldsymbol{U}^{(2)}$ and $\boldsymbol{U}^{(3)}$ are not essential. Note that equation (2.3) together with $\lambda' = b$ yields the value $\lambda_{\text{app}}(t) = \big(\frac{32}{315\pi}\big)^2\big(u^*(0,0)\big)^2 t^4$ given in the formulation of the theorem.

Remark 2.4. In our definitions of $\boldsymbol{U}^{(2)}$ and $\boldsymbol{U}^{(3)}$, the decay at $|x| \to +\infty$ of $W(x)$ plays an important role, which is why we need $N \geq 5$. In large dimensions the nonlinearity $|u|^{\frac{4}{N-2}}u$ becomes quite singular at $u = 0$, which causes some technical difficulties. We chose the easiest dimension $N = 5$.

Step 2. We consider a sequence of solutions $\boldsymbol{u}_n(t)$ of (NLW) with the initial data

$$\boldsymbol{u}_n(t_n) = \Big(u^*(t_n) + W_{\lambda_{\text{app}}(t_n)}, \dot{u}^*(t_n) - \frac{\lambda'_{\text{app}}(t_n)}{\lambda_{\text{app}}(t_n)}\Lambda W(t_n)\Big),$$

where $t_n > 0$ and $\lim_{n\to+\infty} t_n = 0$ (an adjustment has to be made because of the exponential instability near $\boldsymbol{W}$, but this is not a major difficulty). Using a suitable orthogonality condition we decompose $\boldsymbol{u}_n(t) = \boldsymbol{\varphi}(t) + \boldsymbol{g}_n(t)$ and the goal is to control the size of $\boldsymbol{g}_n(t)$, uniformly in n, on a time interval $(t_n, t_0]$, $t_0 > 0$.

To this end, we introduce a *mixed energy-virial functional* $H_n(t)$, which is a small perturbation of the energy functional $E(\boldsymbol{\varphi}(t)+\boldsymbol{g}_n(t)) - E(\boldsymbol{\varphi}(t)) - \langle \mathrm{D}E(\boldsymbol{\varphi}(t)), \boldsymbol{g}_n(t)\rangle$. This functional has the following coercivity property:

$$\|\boldsymbol{g}_n(t)\|^2_{\dot{H}^1\times L^2} \lesssim H_n(t) \qquad \text{(modulo the unstable modes)}.$$

Moreover, the fact that $\boldsymbol{\varphi}(t)$ is a refined ansatz can be used to show that for some large constant C_0 we have

$$\|\boldsymbol{g}_n(t)\|_{\dot{H}^1\times L^2} \leq C_0 t^{\frac{9}{2}} \quad \Rightarrow \quad H'_n(t) \leq c_0 \cdot C_0^2 t^8,$$

with a small constant c_0. A classical continuity argument yields the uniform control $\|\boldsymbol{g}\|_{\dot{H}^1\times L^2} \leq C_0 t^{\frac{9}{2}}$. From this estimate we deduce by standard ODE techniques that the modulation parameter $\lambda_n(t)$ is close to $\lambda_{\text{app}}(t)$.

The conclusion of the second step is that

$$\|\boldsymbol{u}_n(t) - (u^*(t) + W_{\lambda_{\text{app}}}(t), \dot{u}^*(t))\|_{\dot{H}^1\times L^2} \lesssim t^3,$$

uniformly with respect to n.

Step 3. Using the profile decomposition of Bahouri and Gérard [3], we prove a version of sequential weak continuity of the flow, which allows to obtain our solution $\boldsymbol{u}(t)$ as a weak limit of a subsequence of the sequence $\boldsymbol{u}_n(t)$ in the energy space.

The idea of constructing a uniformly controlled sequence of solutions converging to a singular solution was introduced by Merle [64]. Combining this technique with energy estimates was an idea of Martel [56]. Raphaël and Szeftel [78] used a virial correction of the energy functional in a similar context in their study of minimal mass blow-up solutions for the nonlinear Schrödinger equation. The first step of the proof is also inspired by the work of Martel, Merle and Raphaël [60] on exotic blow-up for the L^2-critical generalized Korteweg-de Vries equation. They observed that blow-up rate is directly related to the size of interaction of the bubble with the "background", which is the heart of our construction of the approximate solution.

In fact, the size of this interaction appears explicitly in (2.3), which will be explained in Chapter 1. We find that the main advantage of our method is to demonstrate a direct link between the asymptotic behavior of $\boldsymbol{u}^*(t)$ at $x = 0$ and the asymptotics of $\lambda(t)$. For example, in Theorem 2.2 we have $\dot{u}_0^* \equiv 0$ and $u_0^*(x) \sim |x|^{\frac{\nu-3}{2}}$ in a neighborhood of $x = 0$.

A delicate point, left open by our approach, is to understand better the regularity of the constructed solutions. We have no information on the regularity of the solution besides the fact that it belongs to the energy space. At the same time, it is expected that these solutions have singularities, and working at the energy level allows to avoid dealing with them directly. Note also that using "backward energy estimates" seems to be a natural method of constructing unstable objects. Energy estimates "in the direction of the blow-up" typically yield some sort of stability (modulo a finite codimension), which we do not expect here.

In later works, we hope to carry out a similar construction for equation (WM) and in the case of a singular $\boldsymbol{u}_0^*$.

Remark 2.5. Because of a link with Chapter 3, we would like to point out that the rate of the blow-up can also be formally predicted by solving for $\lambda'(t)$ the equation

$$E\Big(\Big(u^*(t) + W_{\lambda(t)}, \dot u^*(t) - \frac{\lambda'(t)}{\lambda(t)}\Lambda W_{\lambda(t)}\Big)\Big) = E(\boldsymbol{W}) + E(\boldsymbol{u}_0^*).$$

Remark 2.6. Some of the notation in Chapter 1 differ from the ones used above. We preferred to use coherent notation in all the sections of this introductory chapter.

2.2 Construction of two-bubble solutions

In Chapter 2 we construct a radially symmetric solution of (NLW) which exists globally for negative times and decomposes into more that one bubble of energy inside the backward light cone. To the author's knowledge, this is the first example of a solution of this kind.

Theorem 2.7. *There exists a solution* $\boldsymbol{u} : (-\infty, T_0] \to \dot H^1 \times L^2$ *of* (NLW) *in dimension* $N = 6$ *such that*

$$\lim_{t\to-\infty} \|\boldsymbol{u}(t) - (\boldsymbol{W} + \boldsymbol{W}_{\frac{1}{\kappa}\mathrm{e}^{-\kappa|t|}})\|_{\dot H^1\times L^2} = 0, \qquad \text{with } \kappa := \sqrt{\frac{5}{4}}.$$

Note that $\boldsymbol{u}(t)$ is an example of a solution of (NLW) other than the ground state which is global in one time direction and does not contain a radiation term in the energy space, more precisely

$$\forall A > 0,\ \limsup_{t\to-\infty}\int_{|x|\geq|t|-A} |\dot u(t,x)|^2 + |\nabla u(t,x)|^2\,\mathrm{d}x = 0.$$

If we want to emphasize the fact that the energy of our solutions is exactly equal to twice the energy of $\boldsymbol{W}$, with no energy radiated as a free wave, we say that they are *pure two-bubbles*.

Let us say a few words about the proof. Observe that in Theorem 2.1 we can take $\boldsymbol{u}_0^* = (W, 0)$, which produces a blow-up in finite time with $(W, 0)$ as the asymptotic profile. This is "almost" what we desire, except that we need a concentration of the bubble in infinite time rather than a blow-up. A natural way to achieve this is to increase the dimension of the space. In dimension $N = 6$ the interaction between the two bubbles is weaker, which produces the required effect. For $N \geq 7$ one should be able to obtain a similar conclusion, with the speed of concentration of the bubble $\lambda \sim |t|^{-\frac{4}{N-6}}$.

The main technical differences with respect to Theorem 2.1 come from the fact that $\boldsymbol{W}$ has a fixed size in the energy space, whereas before, using the finite speed of propagation, it could be assumed that $\|\boldsymbol{u}_0^*\|_{\dot H^1\times L^2}$ is small. In order to obtain the bounds of the energy norm, we need to study the coercivity of the energy functional in a neighborhood of a sum of two bubbles.

We also extend our construction to the context of equation (WM) with $k \geq 3$. We obtain the following result.

Theorem 2.8. *Fix $k > 2$. There exists a solution $\boldsymbol{u} : (-\infty, T_0] \to \mathcal{E}$ of* (WM) *such that*

$$\lim_{t\to-\infty} \|\boldsymbol{u}(t) - (-\boldsymbol{W} + \boldsymbol{W}_{\frac{k-2}{\kappa}(\kappa|t|)^{-\frac{2}{k-2}}})\|_{\mathcal{E}} = 0, \qquad \text{with } \kappa := \frac{k-2}{2}\Big(\frac{8k}{\pi}\sin\big(\frac{\pi}{k}\big)\Big)^{\frac{1}{k}}.$$

Here, $\boldsymbol{W}$ denotes the ground state of topological degree k and $\mathcal{E}$ is the natural energy space.

Finally, we prove a similar result for the critical Yang-Mills equation:

$$\partial_t^2 u(t,r) = \partial_r^2 u(t,r) + \frac{1}{r}\partial_r u(t,r) - \frac{4}{2r^2}u(t,r)(1-u(t,r))\big(1-\frac{1}{2}u(t,r)\big).$$

Classification results proving that decompositions of type (1.10) and (1.9) hold for a sequence of times $t_n \to T_+$ were obtained by Côte [14] for equation (WM) with $k = 1$, and by Jia and Kenig [42] in greater generality, including all the cases considered here.

2.3 Bounds on the speed of type II blow-up

The results of Chapter 1 encourage to further investigate the relationship between the dynamics of the blow-up and the properties of the asymptotic profile $\boldsymbol{u}_0^*$. In Chapter 3, we prove the following two results in this direction.

Theorem 2.9. *Let $N \in \{3,4,5\}$ and $s > \frac{N-2}{2}$, $s \geq 1$. Let $\boldsymbol{u}_0^* = (u_0^*, \dot u_0^*) \in H^{s+1} \times H^s$ be radially symmetric. Suppose that $\boldsymbol{u}(t)$ is a radial solution of* (NLW) *such that*

$$\lim_{t\to T_+} \|\boldsymbol{u}(t) - \boldsymbol{W}_{\lambda(t)} - \boldsymbol{u}_0^*\|_{\dot H^1\times L^2} = 0, \qquad \lim_{t\to T_+}\lambda(t) = 0, \qquad T_+ < +\infty. \tag{2.4}$$

There exists a constant $C > 0$ depending on $\boldsymbol{u}_0^*$ such that:

- *if $N \in \{4,5\}$, then for $T_+ - t$ sufficiently small there holds*

$$\lambda(t) \leq C(T_+ - t)^{\frac{4}{6-N}}.$$

- *if $N = 3$, then there exists a sequence $t_n \to T_+$ such that*

$$\lambda(t_n) \leq C(T_+ - t_n)^{\frac{4}{6-N}}.$$

Theorem 2.10. *Let $N \in \{3,4,5\}$. Let $\boldsymbol{u}_0^* = (u_0^*, \dot u_0^*) \in H^3 \times H^2$ be radially symmetric and*

$$u_0^*(0) < 0.$$

There exist no radial solutions of (NLW) *such that*

$$\lim_{t\to T_+} \|\boldsymbol{u}(t) - \boldsymbol{W}_{\lambda(t)} - \boldsymbol{u}_0^*\|_{\dot H^1\times L^2} = 0, \qquad \lim_{t\to T_+}\lambda(t) = 0, \qquad T_+ < +\infty. \tag{2.5}$$

The main idea of the proofs, sketched in Paragraph 1.4 of Chapter 3, is to bound the error term by the (energetic) interaction of the bubble with the remainder, which was alluded to in Remark 2.5.

The general picture emerging from Theorems 2.1, 2.2 and 2.9 is that the speed of the blow-up of solutions of Krieger-Schlag-Tataru type is related to the regularity of the asymptotic profile rather than the regularity of the solution itself (although the solutions which blow up faster tend to be more regular, as indicated in [52, 53]). Note that, even if for technical reasons we require additional regularity of $\boldsymbol{u}_0^*$, in (2.4) and (2.5) we require convergence just in the energy space.

A natural question is whether the asymptotic profile uniquely determines the solution which blows up in a neighborhood of $\{\boldsymbol{W}_\lambda\}$. If it was the case, then assigning to each blow-up solution its asymptotic profile and vice versa would yield a classification of the blow-up solutions in a neighborhood of $\{\boldsymbol{W}_\lambda\}$ in the energy space. This general scheme, as pointed out to the author by H. Koch, D. Tataru and the anonymous referee of the article constituting Chapter 3, resembles the classical scattering problem. Of course one should identify the elements of the energy space which can act as the asymptotic profile of some blow-up solution, and the content of Theorem 2.10 is to show that at least certain profiles have to be excluded.

Even if we should probably be rather dubious about the program described above, it raises questions which seem more accessible. For example, does the asymptotic profile determine the speed of the blow-up?

2.4 Nonexistence of two-bubbles with opposite signs

Chapter 4 is devoted to a proof of the following fact.

Theorem 2.11. *Let $N \geq 3$. There exist no radial solutions $\boldsymbol{u} : [t_0, T_+) \to \dot H^1 \times L^2$ of* (NLW) *such that*

$$\lim_{t \to T_+} \|\boldsymbol{u}(t) - \boldsymbol{W}_{\lambda(t)} + \boldsymbol{W}_{\mu(t)}\|_{\dot H^1 \times L^2} = 0$$

and

- *in the case $T_+ < +\infty$, $\lambda(t) \ll \mu(t) \ll T_+ - t$ as $t \to T_+$,*
- *in the case $T_+ = +\infty$, $\lambda(t) \ll \mu(t) \ll t$ as $t \to +\infty$.*

An in the setting of Theorem 2.7, we work here at the energy $2E(W)$, the threshold energy for a formation of a multi-bubble. The main motivation for this work comes from equation (WM), where the two bubbles can either have opposite orientations, in which case the topological degree of the solution is 0, or have the same orientation, which would imply that the topological degree of the solution is $\pm 2k$. This first situation was considered in Theorem 2.8, and we know that a two-bubble can form at least for $k \geq 3$. On the other hand, by well known variational arguments, any k-equivariant map of topological degree $2k$ has energy $> 2E(W)$, hence the conservation of energy implies that there are no pure two-bubbles in this case.

Our proof of Theorem 2.11 is variational in nature, just like the proof for the wave maps described above. The dynamical content is reduced to the study of the hyperbolic dynamics induced by the presence of the linear stable and unstable modes in a neighborhood of $\boldsymbol{W}$, which is an important difference with respect to equation (WM). The most difficult part is to exclude global solutions which behave asymptotically as a superposition of a positive bubble at scale 1 and a negative concentrating bubble. To this end, using the *stable manifold* constructed by Duyckaerts and Merle [27], we introduce *refined* (nonlinear) instability directions, which allow to obtain sharp coercivity properties of the energy functional. The scheme of the proof is influenced by the results of Krieger, Nakanishi and Schlag [49], which we already mentioned in Paragraph 1.4.

Chapter 1

Construction of type II blow-up solutions for the energy-critical wave equation in dimension 5

Abstract

We consider the semilinear wave equation with focusing energy-critical nonlinearity in space dimension $N = 5$
$$\partial_{tt} u = \Delta u + |u|^{4/3} u,$$
with radial data. It is known [25] that a solution $(u, \partial_t u)$ which blows up at $t = 0$ in a neighborhood (in the energy norm) of the family of solitons W_λ, decomposes in the energy space as
$$(u(t), \partial_t u(t)) = (W_{\lambda(t)} + u_0^*, u_1^*) + o(1),$$
where $\lim_{t\to 0} \lambda(t)/t = 0$ and $(u_0^*, u_1^*) \in \dot{H}^1 \times L^2$. We construct a blow-up solution of this type such that the asymptotic profile (u_0^*, u_1^*) is any pair of sufficiently regular functions with $u_0^*(0) > 0$. For these solutions the concentration rate is $\lambda(t) \sim t^4$. We also provide examples of solutions with concentration rate $\lambda(t) \sim t^{\nu+1}$ for $\nu > 8$, related to the behaviour of the asymptotic profile near the origin.

1 Introduction

1.1 General setting

We are interested in the problem of constructing type II blow-up solutions for the energy-critical wave equation in space dimension $N = 5$:

$$\partial_{tt}u = \Delta u + |u|^{4/3}u, \qquad (t, x) \in \mathbb{R} \times \mathbb{R}^5.$$

Denote $f(u) := |u|^{4/3}u$. It will be convenient to write the wave equation as a first-order in time system:

$$\begin{cases} \dfrac{\partial}{\partial t}\begin{pmatrix} u \\ \partial_t u\end{pmatrix} = \begin{pmatrix} \partial_t u \\ \Delta u + f(u)\end{pmatrix}, \\ \begin{pmatrix} u(t_0) \\ \partial_t u(t_0)\end{pmatrix} = \begin{pmatrix} u_0 \\ u_1\end{pmatrix} \in \dot H^1 \times L^2. \end{cases} \tag{NLW}$$

This equation is locally well-posed in the energy space $\dot H^1 \times L^2$ (see for example [47] and the references therein). In particular, for any initial data (u_0, u_1) there exists a maximal interval of existence (T_-, T_+), $-\infty \leq T_- < t_0 < T_+ \leq +\infty$, and a unique solution $(u, \partial_t u) \in C((T_-, T_+); \dot H^1 \times L^2)$. This solution conserves the energy:

$$E(u(t), \partial_t u(t)) = \frac{1}{2}\int |\partial_t u|^2\,\mathrm{d}x + \frac{1}{2}\int |\nabla u|^2\,\mathrm{d}x - \int F(u)\,\mathrm{d}x = E(u_0, u_1),$$

where $F(u) = \int f(u)\,\mathrm{d}u = \frac{3}{10}|u|^{10/3}$ (notice that $\int F(u)\,\mathrm{d}x$ is finite by the Sobolev imbedding).

For a function $v : \mathbb{R}^5 \to \mathbb{R}$ and $\lambda > 0$, we denote

$$v_\lambda(x) := \frac{1}{\lambda^{3/2}}v\big(\frac{x}{\lambda}\big), \qquad v_{\underline{\lambda}}(x) := \frac{1}{\lambda^{5/2}}v\big(\frac{x}{\lambda}\big).$$

A change of variables shows that

$$E\big((u_0)_\lambda, (u_1)_{\underline{\lambda}}\big) = E(u_0, u_1).$$

Equation (NLW) is invariant under the same scaling. If $(u, \partial_t u)$ is a solution of (NLW) and $\lambda > 0$, then

$$t \mapsto \Big(u\big(\frac{t - t_0}{\lambda}\big)_\lambda, \partial_t u\big(\frac{t-t_0}{\lambda}\big)_{\underline{\lambda}}\Big)$$

is also a solution with initial data $\big((u_0)_\lambda, (u_1)_{\underline{\lambda}}\big)$ at time $t = 0$. This is why equation (NLW) is called *energy-critical.*

We introduce also the infinitesimal generators of scale change:

$$\Lambda v := -\partial\lambda v_\lambda\Big|_{\lambda=1} = \big(\frac{3}{2} + x\cdot\nabla\big)v,$$
$$\Lambda_0 v := -\partial\lambda v_{\underline{\lambda}}\Big|_{\lambda=1} = \big(\frac{5}{2} + x\cdot\nabla\big)v.$$

A fundamental object in the study of (NLW) is the family of solutions $(u, \partial_t u) = (W_\lambda, 0)$, where

$$W(x) = \Big(1 + \frac{|x|^2}{15}\Big)^{-3/2}.$$

The functions W_λ are called *ground states.* In this paper we are interested in radial solutions $(u, \partial_t u)$ of (NLW) such that $\inf_\lambda \|(u - W_\lambda, \partial_t u)\|_{\dot H^1\times L^2}$ remains small for $T_- < t \leq t_0$. In the case $N = 3$ it was proved by Krieger, Nakanishi and Schlag [49] that such solutions form a codimension one manifold in a neighbourhood of the family $\{W_\lambda\}$. This is expected to hold also for $N = 5$. The asymptotic behaviour of such (not necessarily radial) solutions as $t \to T_-$ was described by Duyckaerts, Kenig and Merle in [25], both in the case $T_- = -\infty$ and $T_- > -\infty$. In the second case, which is relevant for us, they obtain the following result.

Theorem. *[25, Theorem 2] Let $(u, \partial_t u)$ be a solution of* (NLW) *such that $T_- = 0$ and $\inf_\lambda \|(u - W_\lambda, \partial_t u)\|_{\dot H^1 \times L^2}$ remains small for $T_- < t \le T_0$. Then there exists a C^0 function $\lambda(t) : (0, T_0) \to (0, +\infty)$, such that*

$$\lim_{t \to 0^+} \big(u(t) - W_{\lambda(t)}, \partial_t u(t)\big) = (u_0^*, u_1^*) \in \dot H^1 \times L^2,$$

and the convergence is strong in $\dot H^1 \times L^2$. In addition, $\lambda(t) \ll t$ as $t \to 0^+$.

In this context, W_λ is called the *bubble of energy* and (u_0^*, u_1^*) is called the *asymptotic profile.*

Solutions of this type were first constructed by Krieger, Schlag and Tataru [53] in space dimension $N = 3$, where it is shown that for any $\nu > 1/2$ there exists a solution such that the concentration speed is $\lambda(t) \sim t^{1+\nu}$ (later Krieger and Schlag [50] improved this to $\nu > 0$). Similar results where obtained for energy-critical wave maps by the same authors [52], for energy-critical NLS in dimension $N = 3$ by Ortoleva and Perelman [74] and for energy-critical Schrödinger maps by Perelman [76]. Using a different approach, Hillairet and Raphaël [36] obtained C^∞ blow-up solutions for energy-critical wave equation in dimension $N = 4$ with blow-up rate $\lambda(t) = t \exp\big(-\sqrt{-\log t}(1 + o(1))\big)$. Collot [13] obtained a related result for supercritical wave equation in large dimension.

It follows from the classification of solutions with energy $E(W)$ by Duyckaerts and Merle [27] that necessarily $(u_0^*, u_1^*) \neq 0$. In other words, we have non-existence of minimal energy blow-up solutions. Analogous result is true also for energy-critical wave maps, energy-critical Schrödinger maps and energy-critical NLS.

This is in contrast with the L^2-critical NLS where the conformal invariance produces explicit solutions concentrating a bubble of mass and tending weakly to 0 at blow-up. Existence of blow-up solutions with a non-zero smooth asymptotic profile was first observed by Bourgain and Wang [5]. Blow-up solutions close to the ground state in the case of L^2-critical NLS were extensively studied in a series of papers by Merle and Raphaël. They examined in particular the relationship between regularity of the asymptotic profile and the blow-up speed. One can consult a survey [61] for an account of these results in a proper perspective and a presentation of recent developpements in the case of L^2-critical gKdV.

1.2 Main results

The aim of this paper is to construct solutions which blow up by concentration of one bubble of energy in space dimension $N = 5$. Our approach differs substantially from [53] in that it produces a blow-up solution with a given asymptotic profile. This profile is seen as a source term which permits concentration of the bubble. This point of view is close to a recent construction by Martel, Merle and Raphaël [60] in the case of L^2-critical gKdV.

Denote $X^s := \dot H^{s+1} \cap \dot H^1$. We prove the following two results.

Theorem 1. *Let $(u_0^*, u_1^*) \in X^4 \times H^4$ be any radial functions with $u_0^*(0) > 0$. Let $(u^*(t), \partial_t u^*(t))$ be the solution of* (NLW) *for the initial data $(u^*(0), \partial_t u^*(0)) = (u_0^*, u_1^*)$. There exists a solution $(u, \partial_t u)$ of* (NLW) *defined on a time interval $(0, T_0)$ and a C^1 function $\lambda(t) : (0, T_0) \to$*

$(0,+\infty)$ *such that*

$$\|(u(t)-W_{\lambda(t)}-u^*(t),\ \partial_t u(t)+\lambda'(t)(\Lambda W)_{\underline{\lambda(t)}}-\partial_t u^*(t))\|_{\dot H^1\times L^2}=O(t^{9/2})\qquad \text{as } t\to 0^+, \quad (1.1)$$

and $\lambda(t)=\left(\frac{32}{315\pi}\right)^2\left(u^*(0,0)\right)^2t^4+o(t^4)$.

Theorem 2. *Let* $\nu>8$. *There exists a solution* $(u,\partial_t u)$ *of* (NLW) *defined on the time interval* $(0,T_0)$ *such that*

$$\lim_{t\to 0^+}\left\|\left(u(t)-W_{\lambda(t)}-u_0^*,\ \partial_t u(t)-u_1^*\right)\right\|_{\dot H^1\times L^2}=0,$$

where $\lambda(t)=t^{\nu+1}$, *and* (u_0^*,u_1^*) *is an explicit radial* C^2 *function.*

We will refer to the situation of Theorem 1 as the *non-degenerate case* and to the situation of Theorem 2 as the *degenerate case*. Note that in Theorem 1 we allow any regular (u_0^*,u_1^*) with $u_0^*(0)>0$. Our result might be seen as a first step in a possible classification of all blow-up solutions with a non-degenerate asymptotic profile. Theorem 2 demonstrates how the asymptotic behaviour of (u_0^*,u_1^*) at $x=0$ influences the blow-up speed. The condition $\nu>8$ is imposed by our method. It could be improved at the cost of some technical details, but we are far from obtaining the whole range $\nu>0$ as in [53] for $N=3$.

Let us mention that radiality is only a simplifying assumption. All the estimates used here are true also in the non-radial situation.

In Theorem 2, the function u_0^* is given explicitely by (4.1) and $u_1^*=0$. It follows from our proof that there exists a C^1 function $\widetilde\lambda(t):(0,T_0)\to(0,+\infty)$ such that $\widetilde\lambda(t)=t^{\nu+1}+o(t^{\nu+1})$ and the solution $(u,\partial_t u)$ satisfies

$$\left\|\left(u(t)-W_{\widetilde\lambda(t)}-u^*(t),\ \partial_t u(t)+\widetilde\lambda'(t)(\Lambda W)_{\underline{\widetilde\lambda(t)}}-\partial_t u^*(t)\right)\right\|_{\dot H^1\times L^2}=O(t^{\frac76\nu-\frac43}).$$

1.3 Structure of the proof

In Section 2 we present a formal computation which explains the relation between the asymptotic behaviour of (u_0^*,u_1^*) and the blow-up speed, as well as the relevance of the condition $u_0^*(0)>0$.

In Section 3 we specify an ansatz $(\varphi_0(t),\varphi_1(t))$ in the non-degenerate case and prove appropriate bounds on the error of this approximate solution.

In Section 4 we choose $(u_0^*(0),u_1^*(0))$ such that the same procedure leads to an approximate solution with $\lambda(t)\sim t^{1+\nu}$, and we prove appropriate bounds on the error in this situation.

Section 5 covers both the non-degenerate and the degenerate case. We use a well-known compactness argument introduced by Merle [64] and used by several authors starting with the work of Martel [56] for constructions of multi-solitons. We take a decreasing sequence $t_n\to 0^+$ and we define $(u_n,\partial_t u_n)$ as the solution of (NLW) such that $(u_n(t_n),\partial_t u_n(t_n))$ is close to the approximate solution at time $t=t_n$. The heart of the analysis is to obtain uniform energy bounds for this sequence. That is to say, there exists $T_0>0$ such that $(u_n(t),\partial_t u_n(t))$ stays close to $(\varphi_0(t),\varphi_1(t))$ for $t_n\le t\le T_0$, with bounds independent of n. Note that the exponential instability of W_λ causes an additional difficulty in the argument. We use the shooting method to eliminate the unstable mode. The blow-up solution $(u,\partial_t u)$ is obtained as a weak limit of a subsequence of $(u_n,\partial_t u_n)$. To obtain the crucial uniform energy bounds, we use a mixed energy-virial functional. This method was introduced by Raphaël and Szeftel [78] for a construction of minimal mass blow-up solutions for NLS.

In Appendix A we prove sequential weak continuity of the dynamical system (NLW) under some natural (non-optimal) condition, which is an adaptation of an analogous result of Bahouri and Gérard in the defocusing case [3, Corollary 1]. This result is required in order to extract a weak limit of the sequence $(u_n, \partial_t u_n)$.

In Appendix B we provide for reader's convenience some well-known estimates of the $X^1 \times H^1$ norm of solutions of (NLW). The persistence of $X^1 \times H^1$ regularity is used in Section 5. The energy estimates are used in Section 4. They are non-optimal, but sufficient for our purposes. We prove also propagation of regularity in a neighbourhood of the origin in the non-degenerate case, which is used in Section 3.

1.4 Notation

For $v, w \in L^2$ we denote

$$\langle v, w\rangle := \int_{\mathbb{R}^5} v \cdot w \, \mathrm{d}x.$$

We use the same notation for the duality pairing when $v \in \dot H^{-s}$ and $w \in \dot H^s$.

Linearizing $-\Delta V - f(V)$ around $V = W_\lambda$ we obtain a self-adjoint operator

$$L_\lambda h := -\Delta h - f'(W_\lambda)h.$$

Differentiating $-\Delta W_\lambda - f(W_\lambda) = 0$ with respect to λ we find

$$L_\lambda (\Lambda W)_\lambda = 0.$$

We denote $L := L_1 = -\Delta - f'(W)$.

We will also use the notation $\boldsymbol{v}(t) := (v(t), \partial_t v(t))$.

We denote $\mathcal{Z}$ a fixed radial C_0^∞ function such that $\langle \Lambda W, \mathcal{Z}\rangle > 0$.

Finally, χ is a fixed standard C^∞ cut-off function ($\chi(r) = 1$ for $r \leq 1$, $\chi(r) = 0$ for $r \geq 2$, $\chi'(r) \leq 0$).

2 Formal picture and construction of blow-up profiles

2.1 Inverting the operator L

We define

$$\kappa := -\frac{\langle \Lambda W, f'(W)\rangle}{\langle \Lambda W, \Lambda W\rangle} = \frac{128}{105\pi}.$$

Proposition 2.1. *There exist radial functions $A, B \in C^\infty(\mathbb{R}^5)$ such that*

$$LA = \kappa \Lambda W + f'(W), \qquad LB = -\Lambda_0 \Lambda W. \tag{2.1}$$

In addition, $A(r) \sim r^{-1}$, $A'(r) \sim r^{-2}$, $A''(r) \sim r^{-3}$ and $B(r) \sim r^{-1}$, $B'(r) \sim r^{-2}$, $B''(r) \sim r^{-3}$ as $r \to +\infty$.

Proof. In the proof we will use some standard facts from the theory of Sturm-Liouville equations, see for example [92, Chapter 5].

Solving equation (2.1) is equivalent to solving the following ODE:

$$-(p(r)y')' + q(r)y = g(r), \tag{2.2}$$

with $r \in (0, +\infty)$, $p(r) = r^4$, $q(r) = -r^4 f'(W)$ and $g(r) = g_A(r) = r^4(\kappa \Lambda W(r) + f'(W(r)))$ or $g(r) = g_B(r) = -r^4 \Lambda_0 \Lambda W(r)$. Notice that $|g(r)| \lesssim r^4$ for small r.

We know that $\Lambda W(r)$ is a solution of (2.2) with $g(r) = 0$. Let $\Gamma(r)$ be a second solution normalized in such a way that

$$\mathcal{W}(\Lambda W, \Gamma) = r^4(\Lambda W \cdot \Gamma' - (\Lambda W)' \cdot \Gamma) = 1$$

($\mathcal{W}$ is the modified wronskian, in particular its value is independent of r).

Take $r_1 < \sqrt{15}$, $r_2 > \sqrt{15}$ (recall that $r = \sqrt{15}$ is the unique point where ΛW vanishes) and define

$$y_1(r) := \Lambda W(r) \cdot \int_{r_1}^{r} \frac{\mathrm{d}s}{s^4(\Lambda W(s))^2}, \qquad \text{for } r < \sqrt{15},$$
$$y_2(r) := \Lambda W(r) \cdot \int_{r_2}^{r} \frac{\mathrm{d}s}{s^4(\Lambda W(s))^2}, \qquad \text{for } r > \sqrt{15}.$$

It can be easily checked that y_1 and y_2 are solutions of the homogeneous equation and verify $\mathcal{W}(\Lambda W, y_1) = \mathcal{W}(\Lambda W, y_2) = 1$. Hence, we have $y_j = a_j\Lambda W + \Gamma$ for some scalar coefficients a_1, a_2. Directly from the formulas defining y_1 and y_2 we obtain the asymptotic behaviour of y_1 as $r \to 0^+$ and of y_2 as $r \to \infty$:

$$y_1(r) \sim -\int_{r}^{r_1} \frac{\mathrm{d}s}{s^4} \sim -\frac{1}{r^3}, \qquad r \to 0^+,$$
$$y_2(r) \sim \frac{-1}{r^3}\int_{r_2}^{r} \frac{\mathrm{d}s}{s^4 \cdot s^{-6}} \sim -1, \qquad r \to +\infty.$$

As adding a constant multiple of ΛW does not change these asymptotics, we obtain that $\Gamma(r) \sim -r^{-3}$ as $r \to 0^+$ and $\Gamma(r) \sim -1$ as $r \to +\infty$. From the relation $\mathcal{W}(\Lambda W, \Gamma) = 1$ we get

$$\Gamma' = \frac{r^{-4} + (\Lambda W)' \cdot \Gamma}{\Lambda W},$$

which immediately gives $\Gamma'(r) \sim r^{-4}$ as $r \to 0$ and $\Gamma'(r) \sim \pm r^{-1}$ as $r \to +\infty$ (it can be checked that the sign is " + ", but we will not use this fact).

For $r_0, r \in (0, +\infty)$ we define

$$s(r, r_0) := \Lambda W(r_0)\Gamma(r) - \Gamma(r_0)\Lambda W(r). \tag{2.3}$$

We see that $s(r_0, r_0) = 0$ and $r_0^4 \frac{\mathrm{d}}{\mathrm{d}r}s(r, r_0)|_{r=r_0} = 1$, which means that $s(r, r_0)$ is the second fundamental solution of (2.2). Now using the Duhamel formula we obtain a solution of the non-homogeneous equation (2.2):

$$A(r) = \int_0^r s(r, r')g_A(r')\,\mathrm{d}r',$$
$$B(r) = \int_0^r s(r, r')g_B(r')\,\mathrm{d}r'.$$

Fix $r > 0$ and let $|h| \leq \frac{1}{2}r$. In the estimates which follow, all the constants may depend on r. We have

$$\Big|\frac{A(r+h) - A(r)}{h} - \int_0^r \frac{\mathrm{d}}{\mathrm{d}r}s(r, r')g_A(r')\,\mathrm{d}r'\Big|$$
$$\leq \int_0^r \Big|\frac{s(r+h, r') - s(r, r')}{h} - \frac{\mathrm{d}}{\mathrm{d}r}s(r, r')\Big| \cdot |g_A(r')|\,\mathrm{d}r'$$
$$+ \frac{1}{h}\int_r^{r+h} |s(r+h, r')| \cdot |g_A(r')|\,\mathrm{d}r'.$$

Formula (2.3) implies that $|s(\widetilde{r}, r_0)| \lesssim h$ when $|\widetilde{r} - r| \leq h$ and $|r - r_0| \leq h$. Hence, the second term above converges to 0 as $h \to 0$. For $0 \leq r_0 \leq r$ and $|\widetilde{r} - r| \leq \frac{1}{2}r$ we have the bound $|\frac{\mathrm{d}^2}{\mathrm{d}r^2} s(\widetilde{r}, r_0)| \lesssim r_0^{-3}$. This implies

$$\Big|\frac{s(r+h, r') - s(r, r')}{h} - \frac{\mathrm{d}}{\mathrm{d}r} s(r, r')\Big| \leq \frac{1}{2} \sup_{|\widetilde{r} - r| \leq h} \Big|\frac{\mathrm{d}^2}{\mathrm{d}r^2} s(\widetilde{r}, r')\Big| \cdot |h| \lesssim (r')^{-3} \cdot |h|,$$

so the first term above also converges to 0 as $h \to 0$. This shows that $A(r)$ (and similarly $B(r)$) is continuously differentiable and

$$A'(r) = \int_0^r \frac{\mathrm{d}}{\mathrm{d}r} s(r, r') g_A(r') \,\mathrm{d}r',$$
$$B'(r) = \int_0^r \frac{\mathrm{d}}{\mathrm{d}r} s(r, r') g_B(r') \,\mathrm{d}r'.$$

It is clear from these formulas that $\lim_{r \to 0^+} A'(r) = \lim_{r \to 0^+} B'(r) = 0$.

It follows from above considerations that A and B, seen as functions on $\mathbb{R}^5$, are C^1, so they are C^∞ by elliptic regularity.

Now we consider the behaviour of $A(r)$ and $B(r)$ as $r \to +\infty$. From the crucial orthogonality relation $\int_0^{+\infty} \Lambda W(r') g_A(r') \,\mathrm{d}r' = 0$ we deduce that

$$\Big|\int_0^r \Lambda W(r') g(r') \,\mathrm{d}r'\Big| = \Big|\int_r^{+\infty} \Lambda W(r') g(r') \,\mathrm{d}r'\Big| \lesssim r^{-1}.$$

From this and the asymptotics of Γ and g_A it follows that $|A(r)| \lesssim r^{-1}$ and similarly $|B(r)| \lesssim r^{-1}$. Using the asymptotics of Γ' we obtain also $|A'(r)| \lesssim r^{-2}$ and $|B'(r)| \lesssim r^{-2}$. The fact that $|A''(r)| \lesssim r^{-3}$ and $|B''(r)| \lesssim r^{-3}$ follows from the differential equation. □

We define A and B as the solutions of (2.1) satisfying the orthogonality condition

$$\int_{\mathbb{R}^5} \mathcal{Z} \cdot A \,\mathrm{d}x = \int_{\mathbb{R}^5} \mathcal{Z} \cdot B \,\mathrm{d}x = 0. \tag{2.4}$$

2.2 Determination of blow-up speeds

Let $u^*(t, x)$ be the solution of (NLW) for initial data $(u^*(0), \partial_t u^*(0)) = (u_0, u_1)$. At a formal level, while computing the interaction of u^* with the soliton, we will treat u^* as a function constant in space and C^2 in time, $u^*(t, x) \simeq v^*(t)$. (In the non-degenerate case we will take $v^*(t) = u^*(t, 0)$ and in the degenerate case $v^*(t) = qt^\beta$, where q and β are appropriate constants.) We will construct a solution which blows up at $t = 0$ and is defined for small positive t. This means that in our situation the caracteristic length λ will increase in time. The usual method of performing a formal analysis of blow-up solutions in the case of the wave equation consists in defining $b := \lambda_t$ and searching a solution in the form of a power series in b. Following this scheme, we write

$$\begin{cases} u = W_\lambda + u^*(t) + b^2 T_\lambda + \text{lot} \\ \partial_t u = -b(\Lambda W)_{\underline{\lambda}} + \partial_t u^* + \text{lot}. \end{cases}$$

Here, the profile T is undetermined, and we search a convenient blow up speed. Neglecting irrelevant terms and replacing $\lambda_t := \frac{\mathrm{d}}{\mathrm{d}t}\lambda(t)$ by b, we compute

$$\partial_{tt} u = -b_t (\Lambda W)_{\underline{\lambda}} + \frac{b^2}{\lambda} (\Lambda_0 \Lambda W)_{\underline{\lambda}} + \partial_{tt} u^* + \text{lot}.$$

On the other hand,

$$\Delta u + f(u) = -\frac{1}{\lambda} b^2 (LT)_{\underline{\lambda}} + f'(W_\lambda) v^* + \Delta u^* + f(u^*) + \text{lot}.$$

We discover that, formally at least, we should have

$$LT = -\Lambda_0 \Lambda W + \frac{\lambda}{b^2}[b_t \Lambda W + v^*(t)\sqrt{\lambda} f'(W)]. \tag{2.5}$$

Proposition 2.1 shows that if

$$b_t = \kappa v^*(t)\lambda^{1/2}, \tag{2.6}$$

then equation (2.5) has a decaying regular solution $T = B + \frac{v^*(t)\lambda^{3/2}}{b^2} A$. We call equation (2.6) together with the equation $\lambda_t = b$ *formal parameter equations*. In the non-degenerate case $v^*(t) = u^*(t, 0)$ is close to $u^*(0, 0)$, so we expect that there exists a solution of the formal parameter equations which is close to

$$(\lambda(t), b(t)) = \Big(\frac{\kappa^2 u^*(0,0)^2}{144} t^4, \frac{\kappa^2 u^*(0,0)^2}{36} t^3\Big). \tag{2.7}$$

This is indeed the case, as follows from our analysis in Section 5.

In the degenerate case we have $v^*(t) = qt^\beta$, and the formal parameter equations have a solution

$$(\lambda(t), b(t)) = (t^{1+\nu}, (1+\nu)t^\nu) \tag{2.8}$$

if we choose $q = \frac{\nu(1+\nu)}{\kappa}$ and $\beta = \frac{\nu - 3}{2}$.

3 Approximate solution in the non-degenerate case

3.1 Bounds on the profile (P_0, P_1)

The functions A and B from the previous section do not belong to the space $\dot{H}^1$. We will place a cut-off at the light cone, that is at distance t from the center. Given modulation parameters $(\lambda(t), b(t))$, we define:

$$P_0(t) := \chi\big(\frac{\cdot}{t}\big)(\lambda(t)^{3/2} v^*(t) A_{\lambda(t)} + b(t)^2 B_{\lambda(t)}). \tag{3.1}$$

Recall that in the non-degenerate case $v^*(t) = u^*(t, 0) \in C^2$ by Proposition B.6 and Schauder estimates.

Remark 3.1. Because of the finite speed of propagation, without loss of generality we can replace (u_0^*, u_1^*) by $\big(\chi\big(\frac{\cdot}{\rho}\big)u_0^*, \chi\big(\frac{\cdot}{\rho}\big)u_1^*\big)$, where ρ is a strictly positive constant to be chosen later. Thus, without loss of generality we can assume that the support of (u_0^*, u_1^*) is contained in a small ball and that $\|(u_0^*, u_1^*)\|_{X^1 \times H^1}$ is small.

Remark 3.2. The fact that the profile (P_0, P_1) is cut at $r = t = t^1$ can be considered as a coincidence. The power of t has been chosen in order to optimize the estimates. This is the only power for which we can obtain the estimate of the error term which has asymptotically the same size as the profile P_0. Also, for this choice, $\|P_1\|_{L^2}$ (the forth term of the asymptotic expansion which will be defined in a moment) is asymptotically the same as $\|P_0\|_{\dot{H}^1}$. However, the angle of the cone has no significance for us.

Remark 3.3. Notice that the orthogonality condition which we choose to define A and B has little significance due to a relatively fast decay of ΛW. We will use the same orthogonality condition as for the error term, as this choice simplifies slightly the computation. Observe that the fact that $\mathcal{Z}$ has compact support implies that if $\lambda(t) \ll t$, then $\int P_0(t)\mathcal{Z}_\lambda \,\mathrm{d}x = 0$ for small t.

In the error estimates which will follow, on the right hand side we will always replace $\lambda(t)$ by t^4 and $b(t)$ by t^3, as this is the regime that we are going to consider later in the bootstrap argument. In this section, all the constants may depend on $\boldsymbol{u}^*$.

Lemma 3.4. *Assume that $\lambda(t) \sim t^4$ and $b(t) \sim t^3$. Then*

$$\|P_0(t)\|_{\dot H^1} \lesssim t^{9/2}. \tag{3.2}$$

Proof. It is sufficient to show that $\|\chi(\frac{\cdot}{t})A_\lambda\|_{\dot H^1}^2 \lesssim t^{-3}$ (the computation for B_λ is the same). We have

$$\begin{aligned}
\|\chi(\tfrac{\cdot}{t})A_\lambda\|_{\dot H^1}^2 &\simeq \int_0^{+\infty}\big((\chi(\tfrac{r}{t})A_\lambda(r))'\big)^2 r^4\,\mathrm{d}r = \int_0^{+\infty}\big((\chi(\tfrac{\lambda r}{t})A(r))'\big)^2 r^4\,\mathrm{d}r \\
&\lesssim \int_0^{+\infty}\big(\chi(\tfrac{\lambda r}{t})A'(r)\big)^2 r^4\,\mathrm{d}r + \int_0^{+\infty}\big(\tfrac{\lambda}{t}\chi'(\tfrac{\lambda r}{t})A(r)\big)^2 r^4\,\mathrm{d}r \\
&\lesssim \int_0^{2t/\lambda} r^4\frac{1}{r^4}\,\mathrm{d}r + \frac{\lambda^2}{t^2}\int_{t/\lambda}^{2t/\lambda} r^4\frac{1}{r^2}\,\mathrm{d}r \lesssim \frac{t}{\lambda} \sim t^{-3}.
\end{aligned}$$

□

Lemma 3.5. *Assume that $\lambda(t) \sim t^4$ and $b(t) \sim t^3$. Then*

$$\|L_\lambda P_0 - \lambda^{3/2}v^*(t)L_\lambda A_\lambda - b^2 L_\lambda B_\lambda\|_{L^2} \lesssim t^{7/2}.$$

Proof. We will do the computation only for the terms with A. The terms with B are asymptotically the same. We need to check that

$$\big\|(1-\chi(\tfrac{r}{t}))f'(W_\lambda)A_\lambda\big\|_{L^2} + \big\|\Delta\big((1-\chi(\tfrac{r}{t}))A_\lambda\big)\big\|_{L^2} \lesssim t^{-5/2}$$

For the first term we have even some margin since

$$\begin{aligned}
\big\|(1-\chi(\tfrac{r}{t}))f'(W_\lambda)A_\lambda\big\|_{L^2} &= \frac{1}{\lambda}\big\|(1-\chi(\tfrac{\lambda r}{t}))f'(W)A\big\|_{L^2} \\
&\lesssim \frac{1}{\lambda}\Big(\int_{t/\lambda}^{+\infty}(r^{-4}r^{-1})^2 r^4\,\mathrm{d}r\Big)^{1/2} \sim \frac{1}{\lambda}\cdot\big(\frac{\lambda}{t}\big)^{5/2} \sim t^{7/2}.
\end{aligned}$$

For the second term, we have a few possibilities. Recall that $\Delta = \partial_{rr} + \frac{4\partial_r}{r}$. Either the laplacian hits directly A:

$$\big\|(1-\chi(\tfrac{r}{t}))\Delta(A_\lambda)\big\|_{L^2} = \frac{1}{\lambda}\big\|(1-\chi(\tfrac{\lambda r}{t}))\Delta A\big\|_{L^2} \lesssim \frac{1}{\lambda}\Big(\int_{t/\lambda}^{+\infty}(r^{-3})^2 r^4\,\mathrm{d}r\Big)^{1/2} \sim \frac{1}{\lambda}\cdot\sqrt{\frac{\lambda}{t}} \sim t^{-5/2},$$

either one derivative hits χ:

$$\frac{1}{t}\big\|\chi'(\tfrac{r}{t})\frac{\mathrm{d}}{\mathrm{d}r}(A_\lambda)\big\|_{L^2} = \frac{1}{t}\big\|\chi'(\tfrac{\lambda r}{t})A'(r)\big\|_{L^2} \lesssim t^{-1}\Big(\int_{t/\lambda}^{2t/\lambda}(r^{-2})^2 r^4\,\mathrm{d}r\Big)^{1/2} \sim t^{-1}\cdot\sqrt{\frac{t}{\lambda}} \sim t^{-5/2},$$

(and analogously the term $\frac{1}{t}\|\chi'(\frac{r}{t})\frac{4}{r}(A_\lambda)\|_{L^2}$), or two derivatives hit χ, and we get

$$\frac{1}{t^2}\|\chi''(\frac{r}{t})A_\lambda\|_{L^2} = \frac{\lambda}{t^2}\|\chi''(\frac{\lambda r}{t})A\|_{L^2} \lesssim \frac{\lambda}{t^2}\Big(\int_{t/\lambda}^{2t/\lambda}(r^{-1})^2 r^4\,\mathrm{d}r\Big)^{1/2} \sim \frac{\lambda}{t^2}\cdot\big(\frac{t}{\lambda}\big)^{3/2} \sim t^{-5/2}.$$

□

We define $P_1(t)$ as a formal time derivative of $P_0(t)$, which means that we replace λ_t by b and b_t by $\kappa v^*(t)\lambda^{1/2}$, see (2.6), and we do not differentiate the cut-off function. Explicitely, set

$$\begin{aligned} P_1(t) = \chi\big(\frac{\cdot}{t}\big)\big[&v^*(t)\big(\frac{3}{2}\lambda^{3/2}bA_{\underline{\lambda}} - \lambda^{3/2}b(\Lambda A)_{\underline{\lambda}}\big) \\ &+ \lambda^{5/2}\partial_t v^*(t)A_{\underline{\lambda}} + 2\kappa v^*(t)\lambda^{3/2}bB_{\underline{\lambda}} - b^3(\Lambda B)_{\underline{\lambda}}\big]. \end{aligned} \tag{3.3}$$

Notice that in the regime (2.7) the coefficient $\lambda^{5/2}$ is smaller than the other coefficients (all of which are, asymptotically, of the same size). However, we prefer to keep the corresponding term in the definition of P_1.

Lemma 3.6. *Assume that $\lambda(t) \sim t^4$ and $b(t) \sim t^3$. Then*

$$\|P_1(t)\|_{L^2} \lesssim t^{9/2} \tag{3.4}$$

Proof. All the terms except for the one mentioned above have the same asymptotics, so we will do the computation only for the first one. It is sufficient to show that $\|\chi(\frac{\cdot}{t})A_{\underline{\lambda}}\|_{L^2}^2 \lesssim t^{-9}$. We have

$$\begin{aligned} \|\chi(\frac{\cdot}{t})A_{\underline{\lambda}}\|_{L^2}^2 &\sim \|\chi(\frac{\lambda r}{t})A(r)\|_{L^2(r^4\mathrm{d}r)}^2 \\ &\lesssim \int_0^{2t/\lambda}(r^{-1})^2 r^4\,\mathrm{d}r \lesssim (\frac{t}{\lambda})^3 \sim t^{-9}. \end{aligned}$$

□

Our ansatz $\boldsymbol{\varphi}(t) = (\varphi_0(t), \varphi_1(t))$ is defined as follows:

$$\begin{cases} \varphi_0(t) = W_{\lambda(t)} + P_0(t) + u^*(t), \\ \varphi_1(t) = -b(t)(\Lambda W)_{\underline{\lambda(t)}} + P_1(t) + \partial_t u^*(t), \end{cases}$$

where P_0 and P_1 are given by (3.1) and (3.3).

The error term $\boldsymbol{\varepsilon}(t) = (\varepsilon_0(t), \varepsilon_1(t))$ is defined by the formula:

$$\begin{cases} u(t) = \varphi_0(t) + \varepsilon_0(t), \\ \partial_t u(t) = \varphi_1(t) + \varepsilon_1(t). \end{cases}$$

We shall impose the orthogonality condition

$$\int \varepsilon_0 \mathcal{Z}_{\underline{\lambda}}\,\mathrm{d}x = 0.$$

Lemma 3.7. *If $\lambda \sim t^4$, $b \sim t^3$ and t is small enough, then*

$$|\lambda_t - b| \le \|\boldsymbol{\varepsilon}\|_{\dot{H}^1\times L^2}. \tag{3.5}$$

Proof. To find the formula for λ_t, first we write

$$\begin{aligned}-b(\Lambda W)_{\underline{\lambda}} + \partial_t u^* + P_1(t) + \varepsilon_1(t) = \partial_t u = -\lambda_t(\Lambda W)_{\underline{\lambda}} + \partial_t u^* + \partial_t P_0(t) + \partial_t \varepsilon_0 \Rightarrow \\ \partial_t \varepsilon_0 = (\lambda_t - b)(\Lambda W)_{\underline{\lambda}} + (P_1 - \partial_t P_0) + \varepsilon_1.\end{aligned}$$

Notice that for small t and $\lambda \sim t^4$ we have

$$\int (P_1(t) - \partial_t P_0(t))\mathcal{Z}_{\underline{\lambda}}\,\mathrm{d}x = (\lambda_t - b)[\lambda^{3/2}v^*(t)\langle \Lambda A, \mathcal{Z}\rangle_{L^2} + b^2\langle \Lambda B, \mathcal{Z}\rangle_{L^2}].$$

This follows from (2.4) and the fact that $\mathrm{supp}(\mathcal{Z}_{\underline{\lambda}})$ is contained in the light cone for small t. This gives

$$\begin{aligned}0 &= \frac{\mathrm{d}}{\mathrm{d}t}\int \varepsilon_0 \mathcal{Z}_{\underline{\lambda}}\,\mathrm{d}x = \int \partial_t\varepsilon_0\mathcal{Z}_{\underline{\lambda}}\,\mathrm{d}x - \lambda_t\int \varepsilon_0\frac{1}{\lambda}(\Lambda_0\mathcal{Z})_{\underline{\lambda}}\,\mathrm{d}x \\ &= \int (\lambda_t - b)\langle \Lambda W, \mathcal{Z}\rangle + (\lambda_t - b)[\lambda^{3/2}v^*(t)(\langle \Lambda A, \mathcal{Z}\rangle_{L^2} + b^2\langle \Lambda B, \mathcal{Z}\rangle_{L^2}] \\ &+ \langle \varepsilon_1, \mathcal{Z}_{\underline{\lambda}}\rangle_{L^2} - \lambda_t\int \varepsilon_0\frac{1}{\lambda}(\Lambda_0\mathcal{Z})_{\underline{\lambda}}\,\mathrm{d}x,\end{aligned}$$

and we obtain

$$(\lambda_t - b)[\langle \Lambda W, \mathcal{Z}\rangle + \lambda^{3/2}v^*(t)\langle \Lambda A, \mathcal{Z}\rangle + b^2\langle \Lambda B, \mathcal{Z}\rangle] = -\langle \varepsilon_1, \mathcal{Z}_{\underline{\lambda}}\rangle + \lambda_t\langle \varepsilon_0, \frac{1}{\lambda}(\Lambda_0\mathcal{Z})_{\underline{\lambda}}\rangle.$$

Rearranging the terms we get

$$\begin{aligned}\lambda_t =& \Big(1 - \frac{\langle \varepsilon_0, \frac{1}{\lambda}(\Lambda_0\mathcal{Z})_{\underline{\lambda}}\rangle}{\langle \Lambda W, \mathcal{Z}\rangle_{L^2} + \lambda^{3/2}v^*(t)\langle \Lambda A, \mathcal{Z}\rangle + b^2\langle \Lambda B, \mathcal{Z}\rangle}\Big)^{-1}\cdot \\ &\cdot\Big(b - \frac{\langle \varepsilon_1, \mathcal{Z}_{\underline{\lambda}}\rangle}{\langle \Lambda W, \mathcal{Z}\rangle + \lambda^{3/2}v^*(t)\langle \Lambda A, \mathcal{Z}\rangle + b^2\langle \Lambda B, \mathcal{Z}\rangle}\Big).\end{aligned} \tag{3.6}$$

For t small enough, (3.5) follows. □

Remark 3.8. To be precise, our rigourous argument goes the other way round – we use (3.6) and (2.6) to *define* the local evolution of the modulation parameters, and then by doing exactly the same computation as above, but in the opposite direction, we find that the orthogonality condition $\langle \varepsilon_0, \frac{1}{\lambda}\mathcal{Z}_{\underline{\lambda}}\rangle_{L^2} = 0$ is preserved if it is verified at the initial time (which will be the case). Notice also that using (2.4) we obtain

$$\langle u - W_\lambda - u^*, \mathcal{Z}_{\underline{\lambda}}\rangle = 0. \tag{3.7}$$

Differentiating this condition we find

$$\lambda_t\big(\langle \Lambda W, \mathcal{Z}\rangle + \langle \varepsilon_0, \frac{1}{\lambda}(\Lambda_0\mathcal{Z})_{\underline{\lambda}}\big) = -\langle \partial_t u - \partial_t u^*, \mathcal{Z}_{\underline{\lambda}}\rangle. \tag{3.8}$$

We need to estimate the error between the formal and the actual time derivative of P_0:

Lemma 3.9. *Assume that $\lambda(t) \sim t^4$ and $b(t) \sim t^3$. Then*

$$\|\partial_t P_0 - P_1\|_{\dot{H}^1} \lesssim \sqrt{t}(t^3 + \|\boldsymbol{\varepsilon}\|_{\dot{H}^1\times L^2}).$$

Proof. The error has two parts – one comes from differentiating in time the cut off function and the other one from $|\lambda_t - b|$.

$$\begin{aligned}\partial_t P_0 - P_1 = &-\frac{r}{t^2}\chi'\big(\frac{r}{t}\big)(\lambda^{3/2}v^*(t)A_\lambda + b^2 B_\lambda)\\ &+\chi\big(\frac{r}{t}\big)(\lambda_t - b)[v^*(t)\big(\frac{3}{2}\lambda^{3/2}A_{\underline{\lambda}} - \lambda^{3/2}(\Lambda A)_{\underline{\lambda}}\big) - b^2(\Lambda B)_{\underline{\lambda}}].\end{aligned}$$

Using Proposition 2.1, we can write:

$$\begin{aligned}&\|\frac{r}{\lambda}\chi'\big(\frac{r}{t}\big)A_\lambda\|_{\dot H^1} = \|r\chi'\big(\frac{\lambda r}{t}\big)A\|_{\dot H^1}\\ &\lesssim \|\chi'\big(\frac{\lambda r}{t}\big)\cdot\frac{1}{r}\|_{L^2} + \frac{\lambda}{t}\|r\chi''\big(\frac{\lambda r}{t}\big)\cdot\frac{1}{r}\|_{L^2} + \|r\chi'\big(\frac{\lambda r}{t}\big)\cdot\frac{1}{r^2}\|_{L^2}\\ &\lesssim \|\chi'\big(\frac{\lambda r}{t}\big)\cdot\frac{1}{t}\|_{L^2} + \frac{\lambda}{t}\|\chi'\big(\frac{\lambda r}{t}\big)\|_{L^2} \sim \big(\frac{t}{\lambda}\big)^{3/2} \sim t^{-9/2}.\end{aligned}$$

The same computation is valid also for A replaced by B. Now we have

$$\|\frac{r}{t^2}\chi'\big(\frac{r}{t}\big)\lambda^{3/2}A_\lambda\|_{\dot H^1} \lesssim \frac{\lambda}{t^2}\lambda^{3/2}\cdot t^{-9/2} \sim t^2 t^6 t^{-9/2} = t^{7/2},$$

and the same for the second term.

The computation for the second line is similar:

$$\begin{aligned}&\|\chi\big(\frac{r}{t}\big)A_\lambda\|_{\dot H^1} = \|\chi\big(\frac{\lambda r}{t}\big)A\|_{\dot H^1}\\ &\lesssim \|\chi\big(\frac{\lambda r}{t}\big)\cdot\frac{1}{r^2}\|_{L^2} + \frac{\lambda}{t}\|\chi\big(\frac{\lambda r}{t}\big)\cdot\frac{1}{r}\|_{L^2} \sim \sqrt{t/\lambda} \sim t^{-3/2}.\end{aligned}$$

Multiplying by $\sqrt{\lambda}(\lambda_t - b)$ and using Lemma 3.7 we get the desired estimate. The last two terms are exactly the same. ☐

Finally, the following estimate allows to stop the asymptotic expansion of the solution at P_1.

Lemma 3.10. *Assume that $\lambda(t) \sim t^4$ and $b(t) \sim t^3$. Then*

$$\|\partial_t P_1\|_{L^2} \lesssim \sqrt{t}(t^3 + \|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}).$$

Proof. Consider first the terms coming from differentiating the cut-off function. Like in the proof of the previous lemma, we have

$$\|\frac{r}{\lambda}\chi'\big(\frac{r}{t}\big)A_{\underline{\lambda}}\|_{L^2} \lesssim \|\chi'\big(\frac{\lambda r}{t}\big)\|_{L^2} \sim \big(\frac{t}{\lambda}\big)^{5/2},$$

which gives

$$\|\frac{r}{t^2}\chi'\big(\frac{r}{t}\big)v^*(t)\lambda^{3/2}bA_{\underline{\lambda}}\|_{L^2} \lesssim \frac{\lambda}{t^2}\lambda^{3/2}b\cdot\big(\frac{t}{\lambda}\big)^{5/2} \sim t^{7/2}.$$

The term $\|\frac{r}{t^2}\chi'\big(\frac{r}{t}\big)\lambda^{5/2}\partial_t v^*(t)A_{\underline{\lambda}}\|_{L^2}$ is even smaller.

Consider now the other terms. They are of one of the following six types:

- $\chi\big(\frac{r}{t}\big)\lambda_t\lambda^{1/2}bT_{\underline{\lambda}}$,
- $\chi\big(\frac{r}{t}\big)\lambda_t\frac{b^3}{\lambda}T_{\underline{\lambda}}$,

- $\chi\big(\frac{r}{t}\big)b_t\lambda^{3/2}T_{\underline{\lambda}}$,
- $\chi\big(\frac{r}{t}\big)b_t b^2 T_{\underline{\lambda}}$,
- $\chi\big(\frac{r}{t}\big)\lambda_t\lambda^{3/2}\mathrm{d}_t v^*(t)T_{\underline{\lambda}}$,
- $\chi\big(\frac{r}{t}\big)\lambda^{5/2}\mathrm{d}_{tt} v^*(t)T_{\underline{\lambda}}$,

where $T \in \{A, B, \Lambda A, \Lambda B, \Lambda_0 A, \Lambda_0 B, \Lambda_0\Lambda A, \Lambda_0\Lambda B\}$. In all the situations T is regular and decays like r^{-1} (see Proposition 2.1), so we can write

$$\|\chi(\frac{r}{t})T_{\underline{\lambda}}\|_{L^2} \lesssim \Big(\int_{t/\lambda}^{2t/\lambda}\big(\frac{1}{r}\big)^2 r^4\,\mathrm{d}r\Big)^{1/2} \lesssim \big(\frac{t}{\lambda}\big)^{3/2} \sim t^{-9/2}.$$

Using the fact that $\lambda \sim t^4$, $b \sim t^3$, $\lambda_t \lesssim b + \|\boldsymbol{\varepsilon}\|$, $b_t \lesssim \sqrt{\lambda}$ and that $v^*(t)$ is C^2 we obtain

$$\lambda_t\lambda^{1/2}b + \lambda_t\frac{b^3}{\lambda} + b_t\lambda^{3/2} + b_t b^2 + \lambda_t\lambda^{3/2}|\mathrm{d}_t v^*| + \lambda^{5/2}|\mathrm{d}_{tt}v^*| \lesssim t^5(t^3 + \|\boldsymbol{\varepsilon}\|),$$

which finishes the proof. □

The last lemma shows that $\boldsymbol{\varphi}$ is "almost constant" after rescaling.

Lemma 3.11. *Let $c_1 > 0$. If T_0 is sufficiently small, then for $t \in (0, T_0]$ there holds*

$$\|\partial_t(\varphi_0)_{1/\lambda}\|_{\dot H^1} \le \frac{c_1}{t}.$$

Proof. By the definition of φ_0 and P_0 we get

$$(\varphi_0)_{1/\lambda} = W + \chi\big(\frac{\lambda\,\cdot}{t}\big)[\lambda^{3/2}A + b^2 B] + (u^*)_{1/\lambda}.$$

The terms with A and B are similar, so we only consider the first one. We observe that $|\frac{\lambda_t}{\lambda}| \lesssim \frac{1}{t}$ for small t, with an explicit numerical constant. Now

$$\partial_t\big(\chi(\frac{\lambda r}{t})\lambda^{3/2}A\big) = \frac{3}{2}\frac{\lambda_t}{\lambda}\chi\big(\frac{\lambda r}{t}\big)\lambda^{3/2}A - \frac{\lambda r}{t^2}\chi'\big(\frac{\lambda r}{t}\big)\lambda^{3/2}A.$$

The size of the first term is acceptable by Lemma 3.4. For the second one, it is sufficient to notice that $|\frac{\lambda r}{t^2}| \le \frac{2}{t}$ on the support of χ. The conclusion follows again from Lemma 3.4. (Notice that we have a large margin for these two terms.)

Next, we have

$$\big\|\partial_t(u^*)_{1/\lambda}\big\|_{\dot H^1} \le \big\|\partial_t u^*\big\|_{\dot H^1} + \frac{\lambda_t}{\lambda}\big\|\Lambda u^*\big\|_{\dot H^1}.$$

By Proposition B.2 the first term is bounded for small t. Choosing ρ small enough (see Remark 3.1), we can guarantee that $\|\Lambda u^*(t)\|_{\dot H^1}$ will stay small for small t, which is exactly what we need. □

3.2 Error of the ansatz

Our next objective is to estimate the error of the approximate solution, defined as

$$\boldsymbol{\psi}(t) = \begin{pmatrix}\psi_0(t)\\ \psi_1(t)\end{pmatrix} := \begin{pmatrix}\partial_t\varphi_0(t)\\ \partial_t\varphi_1(t)\end{pmatrix} - \begin{pmatrix}\varphi_1(t)\\ \Delta\varphi_0(t) + f(\varphi_0(t))\end{pmatrix}.$$

In order to do this we first need to extract the principal terms of the nonlinear term, which is based on the following pointwise estimate:

Lemma 3.12.

$$|f(k+l+m) - [f(k)+f(m)+f'(k)l+f'(k)m]| \lesssim |f(l)|+f'(l)|k|+f'(m)|k|+f'(m)|l|. \quad (3.9)$$

Proof. The inequality is homogeneous, so we can suppose that $k^2 + l^2 + m^2 = 1$. The right hand side vanishes only for $(k,l,m) \in \{(\pm 1,0,0),(0,0,\pm 1)\}$, so it suffices to prove the inequality in a neighborhood of these 4 points, where it is an easy consequence of the Taylor expansion of f. □

Lemma 3.13. *If $\lambda(t) \sim t^4$, $b \sim t^3$ and t is small, then*

$$\|f(\varphi_0(t)) - [f(W_{\lambda(t)}) + f(u^*) + f'(W_{\lambda(t)})P_0(t) + f'(W_{\lambda(t)})u^*(t)]\|_{L^2} \lesssim t^4.$$

Proof. We put in the preceding lemma $k = W_{\lambda(t)}$, $l = P_0(t)$, $m = u^*(t)$, and we estimate the L^2 norm of the 4 terms on the right hand side of (3.9). When $P_0(t)$ appears, we split it into two parts. We sometimes forget χ, as its presence here can only help (there are no derivatives).

Term "$|f(l)|$":

$$\big(\chi(\frac{r}{t})\lambda^{3/2}A_\lambda\big)^{7/3} \lesssim \chi(\frac{r}{t})\cdot\big(\frac{r}{\lambda}\big)^{-7/3},$$

and $r^{-14/3}$ is integrable near 0, so $\|\big(\big(\chi(\frac{r}{t})\lambda^{3/2}A_\lambda\big)^{7/3}\|_{L^2} \lesssim \lambda^{7/3}\|\chi(\frac{r}{t})r^{-7/3}\|_{L^2} \ll t^4$. In a similar way, $\|\big(\chi(\frac{r}{t})b^2B_\lambda\big)^{7/3}\|_{L^2} \ll t^4$.

Term "$f'(l)|k|$": By a change of variables we get

$$\|(\lambda^{3/2}A_\lambda)^{4/3}W_\lambda\|_{L^2} = \lambda\|A^{4/3}W\|_{L^2} \sim t^4$$

(exponent of λ on the left $= (3/2-3/2)\cdot(4/3) - 3/2 = -3/2$, and the L^2 scaling is $-5/2$).

In a similar way,

$$\|(b^2B_\lambda)^{4/3}W_\lambda\|_{L^2} = \lambda^{-1}b^{8/3}\|B^{4/3}W\|_{L^2} \sim t^4.$$

Term "$f'(m)|k|$": We use once again the L^∞ bound of u^* and the fact that $\|W_\lambda\|_{L^2} \sim \lambda$.

Term "$f'(m)|l|$": Using (3.2) and the fact that $u^*(t)$ is bounded in $L^{20/3}$ for small t (by Proposition B.2), we have

$$\|f'(u^*)P_0\|_{L^2} \leq \|f'(u^*)\|_{L^5}\cdot\|P_0\|_{L^{10/3}} \lesssim t^{9/2}.$$

□

We can now estimate $\boldsymbol{\psi}(t)$.

Proposition 3.14. *Assume that* $\lambda(t) \sim t^4$ *and* $b(t) \sim t^3$. *Then*

$$\|\psi_0(t) + (\lambda_t - b)\frac{1}{\lambda}(\Lambda W)_\lambda\|_{\dot H^1} \lesssim \sqrt{t}(\|\varepsilon(t)\|_{\dot H^1\times L^2} + t^3), \tag{3.10}$$

$$\|\psi_1(t) - (\lambda_t - b)\frac{b}{\lambda}(\Lambda_0\Lambda W)_{\underline{\lambda}}\|_{L^2} \lesssim \sqrt{t}(\|\varepsilon(t)\|_{\dot H^1\times L^2} + t^3). \tag{3.11}$$

Proof. The first inequality is just a reformulation of Lemma 3.9.

For the second inequality, we divide the error into several parts:

$$\begin{aligned}\psi_1 &= \partial t\varphi_1 - (\Delta\varphi_0 + f(\varphi_0))\\ &= (-b_t(\Lambda W)_{\underline{\lambda}} + \frac{b\lambda_t}{\lambda}(\Lambda_0\Lambda W)_{\underline{\lambda}} + \partial t P_1 + \partial_{tt}u^*)\\ &- (\Delta W_\lambda + \Delta P_0 + \Delta u^*)\\ &- (f(W_\lambda) + f(u^*) + f'(W_\lambda)P_0(t) + f'(W_\lambda)u^*),\end{aligned}$$

where we have used Lemma 3.13 in order to raplace $f(\varphi_0)$ by the sum of its principal terms. Rearranging the terms and using (2.6), we can rewrite the sum above as follows:

$$\begin{aligned}\psi_1 &= (\lambda_t - b)\frac{b}{\lambda}(\Lambda_0\Lambda W)_{\underline{\lambda}}\\ &- (\Delta W_\lambda + f(W_\lambda)) + (\partial_{tt}u^* - \Delta u^* - f(u^*))\\ &- v^*(t)\sqrt{\lambda}(-LA + \kappa\Lambda W + f'(W))_{\underline{\lambda}} + \frac{b^2}{\lambda}(LB + \Lambda_0\Lambda W)_{\underline{\lambda}}\\ &+ (-\Delta P_0 - f'(W_\lambda)P_0) - v^*(t)\sqrt{\lambda}(LA)_{\underline{\lambda}} - \frac{b^2}{\lambda}(LB)_{\underline{\lambda}}\\ &+ (v^*(t) - u^*(t))\sqrt{\lambda}(f'(W))_{\underline{\lambda}}\\ &+ \partial_t P_1 + O(t^{7/2}).\end{aligned}$$

Now we proceed line by line.

Line 1. This is the correction that we substract in (3.11).

Line 2. Both terms equal 0.

Line 3. Both terms equal 0 by the definition of A and B.

Line 4. This error is due to the presence of the cut-off function in (3.1), and Lemma 3.5 tells us that it is acceptable.

Line 5. This error is due to the fact that we replace the interaction with $u^*(t)$ by the interaction with the constant in space function $v^*(t)$. It follows from Proposition B.6 that $|v^*(t) - u^*(t,r)| \lesssim r$ uniformly in time when $r \le t$ and t is small. Hence,

$$\|(v^*(t) - u^*(t,r))f'(W_\lambda)\|_{L^2(r\le t)} \lesssim \|r\sqrt{\lambda}(f'(W))_{\underline{\lambda}}\|_{L^2} \sim \lambda^{3/2} \sim t^6.$$

(We have used the fact that $rf'(W) \in L^2$.) In the zone $r \ge t$ first we use the fact that v^* is bounded and

$$\|f'(W_\lambda)\|_{L^2(r\ge t)} = \sqrt{\lambda}\|f'(W)\|_{L^2(r\ge t\lambda^{-1})} \lesssim \sqrt{\lambda}(\lambda/t)^{3/2} \sim t^{13/2}.$$

As for u^*, we know from Proposition B.2 that it is bounded in L^{10}. By Hölder $\|u^* \cdot f'(W_\lambda)\|_{L^2(r\ge t)} \le \|u^*\|_{L^{10}} \cdot \|f'(W_\lambda)\|_{L^{5/2}(r\ge t)}$, and a routine computation shows that the last term is bounded by $(\lambda/t)^2 \sim t^6$.

Line 6. This error is small by Lemma 3.10. □

4 Approximate solution in the degenerate case

4.1 Bounds on the profile (P_0, P_1)

This section is very similar to the previous one. Formula (3.1) is still valid, but recall that in the present case we take $v^*(t) = qt^\beta$ where $q = \frac{\nu(1+\nu)}{\kappa}$ and $\beta = \frac{\nu-3}{2}$. The function u_0^* is defined as follows:

$$u_0^*(x) := \chi\big(\frac{\cdot}{\rho}\big)\cdot p|x|^\beta, \qquad p = \frac{3q}{(\beta+1)(\beta+3)},\ \rho > 0 \text{ small.} \tag{4.1}$$

(by the finite speed of propagation the cut-off does not affect the behaviour at zero for small times, cf. Remark 3.1). We take $u_1^* = 0$.

In the error estimates which will follow, on the right hand side we will always replace $\lambda(t)$ by $t^{1+\nu}$ and $b(t)$ by t^ν, since this is the regime considered later in the bootstrap argument.

Lemma 4.1. *Assume that $\lambda(t) \sim t^{1+\nu}$ and $b(t) \sim t^\nu$. Then*

$$\|P_0(t)\|_{\dot H^1} \lesssim t^{3\nu/2}. \tag{4.2}$$

Proof. Recall that $v^*(t) \sim t^\beta = t^{(\nu-3)/2}$, so $\lambda^{3/2}v^*(t) \sim b^2 \sim t^{2\nu}$. Hence, it is sufficient to show that $\|\chi(\frac{\cdot}{t})A_\lambda\|_{\dot H^1}^2 + \|\chi(\frac{\cdot}{t})B_\lambda\|_{\dot H^1}^2 \lesssim t^{-\nu}$. The computation in the proof of Lemma 3.4 gives

$$\|\chi\big(\frac{\cdot}{t}\big)A_\lambda\|_{\dot H^1}^2 \lesssim \frac{t}{\lambda} \sim t^{-\nu},$$

and similarly for the second term. □

Lemma 4.2. *Assume that $\lambda(t) \sim t^{1+\nu}$ and $b(t) \sim t^\nu$. Then*

$$\|L_\lambda P_0 - \lambda^{3/2}v^*(t)L_\lambda A_\lambda - b^2 L_\lambda B_\lambda\|_{L^2} \lesssim t^{3\nu/2-1}.$$

Proof. We will do the computation only for the terms with A. The terms with B are asymptotically the same. We need to check that

$$\big\|\big(1-\chi\big(\frac{r}{t}\big)\big)f'(W_\lambda)A_\lambda\big\|_{L^2} + \big\|\Delta\big(\big(1-\chi\big(\frac{r}{t}\big)\big)A_\lambda\big)\big\|_{L^2} \lesssim t^{-\nu/2-1}$$

The computations in the proof of Lemma 3.5 imply that the first term is bounded by $\frac{1}{\lambda}\big(\frac{\lambda}{t}\big)^{5/2} \sim t^{3\nu/2-1}$, and the second by

$$\frac{1}{\lambda}\cdot\sqrt{\frac{\lambda}{t}} + \frac{1}{t}\cdot\sqrt{\frac{t}{\lambda}} + \frac{\lambda}{t^2}\cdot\big(\frac{t}{\lambda}\big)^{3/2} \sim (t\cdot\lambda)^{-1/2} \sim t^{-\nu/2-1}.$$

□

In the degenerate case the profile $P_1(t)$ is defined by the same formula (3.3).

Lemma 4.3. *Assume that $\lambda(t) \sim t^{1+\nu}$ and $b(t) \sim t^\nu$. Then*

$$\|P_1(t)\|_{L^2} \lesssim t^{3\nu/2} \tag{4.3}$$

Proof. Notice that $\frac{\mathrm{d}}{\mathrm{d}t}v^*(t) \sim t^{\beta-1} \sim t^{(\nu-5)/2}$. This implies that

$$v^*(t)\cdot\lambda^{3/2}b \sim \frac{\mathrm{d}}{\mathrm{d}t}v^*(t)\cdot\lambda^{5/2} \sim b^3 \sim t^{3\nu},$$

so all the terms in the definition of $P_1(t)$ have asymptotically the same size and it suffices to show that $\|\chi(\frac{\cdot}{t})A_{\underline{\lambda}}\|_{L^2}^2 \lesssim t^{-3\nu}$ (the other terms are similar). The computation in the proof of Lemma 3.6 gives

$$\|\chi(\frac{\cdot}{t})A_{\underline{\lambda}}\|_{L^2}^2 \lesssim (\frac{t}{\lambda})^3 \sim t^{-3\nu}.$$

□

Estimate (3.5) and its proof are valid in the degenerate case.

Lemma 4.4. *Assume that $\lambda(t) \sim t^{1+\nu}$ and $b(t) \sim t^\nu$. Then*

$$\|\partial_t P_0 - P_1\|_{\dot{H}^1} \lesssim t^{\nu/2-1}(t^\nu + \|\boldsymbol{\varepsilon}\|_{\dot{H}^1\times L^2}).$$

Proof. As in the proof of Lemma 3.9, we write

$$\begin{aligned}\partial_t P_0 - P_1 = &-\frac{r}{t^2}\chi'(\frac{r}{t})(\lambda^{3/2}v^*(t)A_\lambda + b^2 B_\lambda)\\ &+\chi(\frac{r}{t})(\lambda_t - b)[v^*(t)(\frac{3}{2}\lambda^{3/2}A_{\underline{\lambda}} - \lambda^{3/2}(\Lambda A)_{\underline{\lambda}}) - b^2(\Lambda B)_{\underline{\lambda}}].\end{aligned}$$

The computation in the proof of Lemma 3.9 implies

$$\|\frac{r}{\lambda}\chi'(\frac{r}{t})A_\lambda\|_{\dot{H}^1} \lesssim (\frac{t}{\lambda})^{3/2} \sim t^{-3\nu/2}.$$

Multiplying by $\frac{\lambda}{t^2}\lambda^{3/2}v^*(t) \sim t^{3\nu-1}$ we obtain the required bound on the first term. The second term of the first line is similar.

The second line is bounded exactly as in the proof of Lemma 3.9. □

Lemma 4.5. *Assume that $\lambda(t) \sim t^{1+\nu}$ and $b(t) \sim t^\nu$. Then*

$$\|\partial_t P_1\|_{L^2} \lesssim t^{\nu/2-1}(t^\nu + \|\boldsymbol{\varepsilon}\|_{\dot{H}^1\times L^2}).$$

Proof. We indicate only the modifications with respect to the proof of Lemma 3.10. The term coming from differentiating the cut-off function is estimated as before by

$$\frac{\lambda}{t^2}v^*(t)\lambda^{3/2}b\cdot(\frac{t}{\lambda})^{5/2} \sim t^{3\nu/2-1}.$$

For the other terms, we get

$$\|\chi(\frac{r}{t})T_{\underline{\lambda}}\|_{L^2} \lesssim (\frac{t}{\lambda})^{3/2} \sim t^{-3\nu/2}.$$

□

4.2 Error of the ansatz

This subsection differs from the non-degenerate case, because we work here only with X^1 regularity and some more effort is required in order to estimate the terms involving u^*.

Lemma 4.6. *If $\lambda(t) \sim t^{1+\nu}$, $b \sim t^\nu$, $\nu > 8$ and t is small, then*

$$\|f(\varphi_0(t)) - [f(W_{\lambda(t)}) + f(u^*) + f'(W_{\lambda(t)})P_0(t) + f'(W_{\lambda(t)})u^*(t)]\|_{L^2} \ll t^{\frac{7}{6}\nu - \frac{7}{3}}.$$

Proof. As in the proof of Lemma 3.13 we use Lemma 3.12 with $k = W_{\lambda(t)}$, $l = P_0(t)$ and $m = u^*(t)$. We obtain that the L^2 norm of the term "$|f(l)|$" is bounded by

$$\big((v^* \cdot \lambda)^{7/3} + b^{14/3}\big)\|\chi(\frac{r}{t})r^{-7/3}\|_{L^2},$$

which is better than required. For the term "$f'(l)|k|$" we obtain the bound $(v^*)^{4/3} \cdot \lambda + b^{4/3}\lambda^{-1} \sim t^{5\nu/3-1}$, which is again better than required.

Term "$f'(m)|k|$": Let $(u_{\text{LIN}}^*, \partial_t u_{\text{LIN}}^*)$ be the solution of the free wave equation for the initial data $(u_{\text{LIN}}^*(0), \partial_t u_{\text{LIN}}^*(0)) = (u_0^*, u_1^*)$. We write

$$\|f'(u^*)\cdot W_\lambda\|_{L^2} \lesssim \|f'(u_{\text{LIN}}^*)\cdot W_\lambda\|_{L^2(|x|\leq \frac{1}{2}t)} + \|f'(u^* - u_{\text{LIN}}^*)\cdot W_\lambda\|_{L^2(|x|\leq\frac{1}{2}t)} + \|f'(u^*)\cdot W_\lambda\|_{L^2(|x|\geq\frac{1}{2}t)}$$

and we examine separately the three terms on the right hand side. It follows from Proposition B.7 that for $|x| \leq \frac{1}{2}t$ we have the bound $|u_{\text{LIN}}^*(t,x)| \lesssim t^\beta = t^{\frac{\nu-3}{2}}$, which implies $\|f'(u_{\text{LIN}}^*(t))\|_{L^\infty} \lesssim t^{\frac{2}{3}(\nu-3)}$, hence

$$\|f'(u_{\text{LIN}}^*)\cdot W_\lambda\|_{L^2(|x|\leq\frac{1}{2}t)} \lesssim t^{\frac{2}{3}(\nu-3)}\|W_\lambda\|_{L^2} \sim t^{\frac{2}{3}(\nu-3)}t^{\nu+1} \ll t^{\frac{7}{6}\nu-\frac{7}{3}}.$$

From Proposition B.8 we infer

$$\|u^* - u_{\text{LIN}}^*\|_{L^{20/3}(|x|\leq\frac{1}{2}t)} \lesssim t^{\frac{7}{6}\nu-\frac{7}{3}},$$

hence

$$\|f'(u^* - u_{\text{LIN}}^*)\|_{L^5(|x|\leq\frac{1}{2}t)} \lesssim t^{\frac{14}{9}\nu-\frac{28}{9}},$$

which leads to

$$\|f'(u^* - u_{\text{LIN}}^*)\cdot W_\lambda\|_{L^2(|x|\leq\frac{1}{2}t)} \leq \|f'(u^* - u_{\text{LIN}}^*)\|_{L^5(|x|\leq\frac{1}{2}t)} \cdot \|W_\lambda\|_{L^{10/3}(|x|\leq\frac{1}{2}t)} \lesssim t^{\frac{14}{9}\nu-\frac{28}{9}},$$

which is more than sufficient for $\nu > 8$.

For $|x| \geq \frac{1}{2}t$, we know from Proposition B.2 that $\|f'(u^*)\|_{L^5}$ is bounded for small t. By a change of variables we obtain

$$\|W_\lambda\|_{L^{10/3}(|x|\geq\frac{1}{2}t)} \lesssim \Big(\int_{t/2\lambda}^{+\infty}(r^{-3})^{10/3}r^4\,\mathrm{d}r\Big)^{\frac{3}{10}} \sim \big(\frac{\lambda}{t}\big)^{3/2} \ll t^{\frac{7}{6}\nu-\frac{7}{3}}.$$

Term "$f'(m)|l|$": Using (4.2) we have

$$\|f'(u^*)\cdot P_0\|_{L^2} \leq \|f'(u^*)\|_{L^5}\cdot\|P_0\|_{\dot{H}^1} \lesssim t^{3\nu/2} \ll t^{\frac{7}{6}\nu-\frac{7}{3}}.$$

□

We can now estimate $\boldsymbol{\psi}(t)$.

Proposition 4.7. *Assume that $\lambda(t) \sim t^{1+\nu}$ and $b(t) \sim t^\nu$. Then*

$$\|\psi_0(t) + (\lambda_t - b)\frac{1}{\lambda}(\Lambda W)_\lambda\|_{\dot H^1} \lesssim t^{\frac{7}{6}\nu - \frac{7}{3}} + t^{\nu/2-1}\|\boldsymbol{\varepsilon}(t)\|_{\dot H^1\times L^2}, \tag{4.4}$$
$$\|\psi_1(t) - (\lambda_t - b)\frac{b}{\lambda}(\Lambda_0\Lambda W)_{\underline{\lambda}}\|_{L^2} \lesssim t^{\frac{7}{6}\nu - \frac{7}{3}} + t^{\nu/2-1}\|\boldsymbol{\varepsilon}(t)\|_{\dot H^1\times L^2}.$$

Proof. The first inequality follows from Lemma 4.4.

For the second inequality, as in the proof of Proposition 3.14, using Lemma 4.6 and rearranging the terms, we get:

$$\begin{aligned}\psi_1 &= (\lambda_t - b)\frac{b}{\lambda}(\Lambda_0\Lambda W)_{\underline{\lambda}}\\
&- (\Delta W_\lambda + f(W_\lambda)) + (\partial_{tt}u^* - \Delta u^* - f(u^*))\\
&- v^*(t)\sqrt{\lambda}(-LA + \kappa\Lambda W + f'(W))_{\underline{\lambda}} + \frac{b^2}{\lambda}(LB + \Lambda_0\Lambda W)_{\underline{\lambda}}\\
&+ (-\Delta P_0 - f'(W_\lambda)P_0) - v^*(t)\sqrt{\lambda}(LA)_{\underline{\lambda}} - \frac{b^2}{\lambda}(LB)_{\underline{\lambda}}\\
&+ (v^*(t) - u^*(t))\sqrt{\lambda}(f'(W))_{\underline{\lambda}}\\
&+ \partial_t P_1 + O(t^{\frac{7}{6}\nu - \frac{7}{3}}).\end{aligned}$$

Lines 1, 2, 3, 4 and 6 are treated exactly as in the proof of Proposition 3.14, using Lemmas 4.2 and 4.5 instead of Lemmas 3.5 and 3.10. We estimate line 5 as follows:

$$\begin{aligned}\|(v^* - u^*)f'(W_\lambda)\|_{L^2} \lesssim\ & \|(v^* - u^*_{\text{LIN}})f'(W_\lambda)\|_{L^2(|x|\le\frac12 t)}\\
&+ \|(u^*_{\text{LIN}} - u^*)f'(W_\lambda)\|_{L^2(|x|\le\frac12 t)}\\
&+ \|v^*\cdot f'(W_\lambda)\|_{L^2(|x|\ge\frac12 t)}\\
&+ \|u^*\cdot f'(W_\lambda)\|_{L^2(|x|\ge\frac12 t)}.\end{aligned}$$

From Proposition B.7 it follows in particular that $|v^*(t) - u^*_{\text{LIN}}(t,r)| \lesssim r$ when $r \le \frac12 t$, hence the proof of Proposition 3.14 gives the bound

$$\|(v^* - u^*_{\text{LIN}})\cdot f'(W_\lambda)\|_{L^2(|x|\le\frac12 t)} \lesssim \lambda^{3/2} \ll t^{\frac{7}{6}\nu - \frac{7}{3}}.$$

From Proposition B.8 and the fact that $\|f'(W_\lambda)\|_{L^{5/2}} = \|f'(W)\|_{L^{5/2}}$ we get

$$\|(u^* - u^*_{\text{LIN}})\cdot f'(W_\lambda)\|_{L^2(|x|\le\frac12 t)} \lesssim t^{\frac73\beta + \frac76} = t^{\frac76\nu - \frac73}.$$

We have

$$\|f'(W_\lambda)\|_{L^2(|x|\ge\frac12 t)} \lesssim \|\frac{\lambda^2}{|x|^4}\|_{L^2(|x|\ge\frac12 t)} \sim \lambda^2 t^{-3/2}$$

and

$$\|f'(W_\lambda)\|_{L^{5/2}(|x|\ge\frac12 t)} \lesssim \|\frac{\lambda^2}{|x|^4}\|_{L^{5/2}(|x|\ge\frac12 t)} \sim \lambda^2 t^{-2}.$$

Using boundedness of v^* in L^∞, boundedness of u^* in L^{10} and Hölder inequality we obtain the required bounds, which terminates the proof. □

Lemma 3.11 is still valid in the degenerate case, as well as its proof (we use Lemma 4.1 instead of Lemma 3.4).

5 Evolution of the error term

The evolution of the error term $\boldsymbol{\varepsilon}$ is governed by the following system of differential equations:

$$\partial_t \begin{pmatrix} \varepsilon_0 \\ \varepsilon_1 \end{pmatrix} = \begin{pmatrix} \varepsilon_1 - \psi_0 \\ \Delta\varepsilon_0 + [f(\varphi_0 + \varepsilon_0) - f(\varphi_0)] - \psi_1. \end{pmatrix}, \tag{5.1}$$

coupled with the equations (3.6) and (2.6) for the modulation parameters $\mathrm{Mod} := (\lambda, b)$. We denote (T_-, T_+) the maximal interval of existence of $\boldsymbol{u}$.

We introduce the energy functional adapted to our ansatz:

$$I(t) := \int \frac{1}{2}|\varepsilon_1|^2 + \frac{1}{2}|\nabla\varepsilon_0|^2 - [F(\varphi_0 + \varepsilon_0) - F(\varphi_0) - f(\varphi_0)\varepsilon_0]\,\mathrm{d}x.$$

Essentially we will perform a bootstrap argument in order to control this functional just by integrating in time its time derivative. We need a virial correction term which is defined as follows:

$$J(t) := b\int \varepsilon_1 \cdot \Big(\frac{1}{\lambda}\cdot\frac{1}{2}(\Delta a)_\lambda + (\nabla a)_\lambda \cdot \nabla\Big)\varepsilon_0\,\mathrm{d}x,$$

where $a_\lambda(r) = a\big(\frac{r}{\lambda}\big)$, $(\nabla a)_\lambda(r) = \nabla a\big(\frac{r}{\lambda}\big)$, $(\Delta a)_\lambda(r) = \Delta a\big(\frac{r}{\lambda}\big)$ and

$$a(r) := \begin{cases} \frac{1}{2}r^2 & |r| \leq R \\ \frac{15}{8}Rr - \frac{5}{2}R^2 + \frac{5}{4}R^3r^{-1} - \frac{1}{8}R^5r^{-3} & |r| \geq R \end{cases}$$

(R is a big radius to be chosen later, see Proposition 5.2).

Lemma 5.1. *The function $a(r)$ defined above, viewed as a function on $\mathbb{R}^5$, has the following properties:*

- $a \in C^{3,1}$,
- *a is strictly convex,*
- $|a(r)| \lesssim r$, $|a'(r)| \lesssim 1$, $|a''(r)| \lesssim r^{-1}$ *when $r \to +\infty$ (the constant depends on R),*
- $-\frac{1}{r^3} \lesssim \Delta^2 a(r) \leq 0$.

Proof. It is apparent from the formula defining a that a is regular except for $r = R$. A computation shows that $a(r)$, $a'(r)$, $a''(r)$ and $a'''(r)$ are Lipschitz near $r = R$. For $r \geq R$ we have $a''(r) = \frac{5}{2}\big(\frac{R}{r}\big)^3 - \frac{3}{2}\big(\frac{R}{r}\big)^5 > 0$, which proves strict convexity. For $r > R$ one can compute $\Delta^2 a(r) = -\frac{15}{r^3}\cdot\frac{1}{R^2}$ (where $\Delta = \partial_{rr} + \frac{4}{r}\partial_r$ is the laplacian in dimension $N = 5$). $\square$

We define the mixed energy-virial functional:

$$H(t) = I(t) + J(t).$$

The proof of the following result, which will occupy most of this section, is valid both in the non-degenerate and the degenerate case. The non-degenerate case is obtained for $\nu = 3$. We denote also:

$$\gamma := \begin{cases} \frac{7}{2} & \text{in the non-degenerate case,} \\ \frac{7}{6}\nu - \frac{7}{3} & \text{in the degenerate case,} \end{cases}$$

which is the exponent of t in the error estimates in Proposition 3.14 and Proposition 4.7 respectively.

We will use the notation:

$$\|\nabla_{a,\lambda}\varepsilon_0\|_{L^2}^2 := \int \sum_{i,j} (\partial_{ij}a)_\lambda \partial_i\varepsilon_0 \partial_j\varepsilon_0 \,\mathrm{d}x.$$

Proposition 5.2. *Let $\nu = 3$ or $\nu > 8$. Suppose that $\lambda \sim t^{1+\nu}$, $b \sim t^\nu$ and let $c > 0$. If R is chosen large enough, then there exist strictly positive constants T_0 and C_1 such that for $[T_1, T_2] \subset (0, T_0] \cap (T_-, T_+)$ there holds*

$$H(T_2) \le H(T_1) + \int_{T_1}^{T_2} \Big(-\frac{b}{\lambda}\Big(\|\nabla_{a,\lambda}\varepsilon_0\|_{L^2}^2 - \int \big(f(\varphi_0+\varepsilon_0) - f(\varphi_0)\big)\varepsilon_0 \,\mathrm{d}x\Big) + \Big(\frac{c}{t}\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}^2 + C_1 t^\gamma \cdot \|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}\Big)\Big)\,\mathrm{d}t. \tag{5.2}$$

The proof of this result is going to be an algebraic computation which is not justified in the space $\dot H^1 \times L^2$. However, we do not need any uniform control of the regularity or the decay, so we can use the following density argument. We can approximate a given $\boldsymbol{\varepsilon}$ in $\dot H^1 \times L^2$ in such a way that the initial data $(u(T_1), \partial_t u(T_1))$ will be in $X^1 \times H^1$ and of compact support. Then locally the evolution will have the same proprieties by Proposition B.5, and will be close to the original one in $\dot H^1 \times L^2$ for all $t \in [T_1, T_2]$ by local well-posedness in $\dot H^1 \times L^2$. The new $\boldsymbol{\varepsilon}$ has sufficient regularity and decay to justify all the computations. Since the estimate (5.2) depends continuously (in $\dot H^1 \times L^2$) on $\boldsymbol{\varepsilon}$, we are done.

We shall split the proof of Proposition 5.2 into several Lemmas. We always work under the hypotheses of Proposition 5.2, that is $\lambda \sim t^{1+\nu}$, $b \sim t^\nu$ and $\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2} \le t^{\gamma+1}$. Notice that $\gamma + 1 > \nu$. In the non-degenerate case $\gamma + 1 = \frac{9}{2} > 3 = \nu$ and in the degenerate case $\gamma + 1 = \frac{7}{6}\nu - \frac{4}{3} > \nu$ because $\nu > 8$. This means that $\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2} \ll b$ and $\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2} \ll \frac{\lambda}{t}$ for small t. In what follows c stands for any small strictly positive constant.

We use the method introduced in [78], which consists in differentiating the nonlinear term in self-similar variables. The resulting error will be corrected by the virial term J. Concretely, we have:

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\int [F(\varphi_0+\varepsilon_0) - F(\varphi_0) - f(\varphi_0)\varepsilon_0]\,\mathrm{d}x \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\int [F((\varphi_0)_{1/\lambda} + (\varepsilon_0)_{1/\lambda}) - F((\varphi_0)_{1/\lambda}) - f((\varphi_0)_{1/\lambda})(\varepsilon_0)_{1/\lambda}]\,\mathrm{d}x \\
&= \int [f((\varphi_0)_{1/\lambda} + (\varepsilon_0)_{1/\lambda}) - f((\varphi_0)_{1/\lambda}) - f'((\varphi_0)_{1/\lambda})(\varepsilon_0)_{1/\lambda}]\partial_t\big((\varphi_0)_{1/\lambda}\big)\,\mathrm{d}x \\
&+ \int [f((\varphi_0)_{1/\lambda} + (\varepsilon_0)_{1/\lambda}) - f((\varphi_0)_{1/\lambda})]\big((\varepsilon_{0t})_{1/\lambda} + \frac{\lambda_t}{\lambda}(\Lambda\varepsilon_0)_{1/\lambda}\big)\,\mathrm{d}x.
\end{aligned}$$

The first term can be neglected, as shown by Lemma 3.11. Scaling back the second term we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\int [F(\varphi_0+\varepsilon_0) - F(\varphi_0) - f(\varphi_0)\varepsilon_0]\,\mathrm{d}x \simeq \int [f(\varphi_0+\varepsilon_0) - f(\varphi_0)]\big(\varepsilon_{0t} + \frac{\lambda_t}{\lambda}\Lambda\varepsilon_0\big)\,\mathrm{d}x. \tag{5.3}$$

Here and later the sign $\simeq$ means that the difference of the two sides has size at most $\frac{c}{t}\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}^2 + C_1 t^\gamma \cdot \|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}$. Also, when we say that a term is "negligible", it always means that its absolute value is bounded by $\frac{c}{t}\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}^2 + C_1 t^\gamma \cdot \|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}$.

Using the equations (5.1), (5.3) and integrating by parts, we obtain standard cancellations:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}I(t) &\simeq \int \varepsilon_1\varepsilon_{1t}\,\mathrm{d}x - \int [\Delta\varepsilon_0 + f(\varphi_0+\varepsilon_0) - f(\varphi_0)]\varepsilon_{0t}\,\mathrm{d}x \\
&\quad - \frac{\lambda_t}{\lambda}\int [f(\varphi_0+\varepsilon_0) - f(\varphi_0)]\Lambda\varepsilon_0\,\mathrm{d}x \\
&\simeq -\int \varepsilon_1\psi_1\,\mathrm{d}x + \int [\Delta\varepsilon_0 + f(\varphi_0+\varepsilon_0) - f(\varphi_0)]\psi_0\,\mathrm{d}x \\
&\quad - \frac{\lambda_t}{\lambda}\int [f(\varphi_0+\varepsilon_0) - f(\varphi_0)]\Lambda\varepsilon_0\,\mathrm{d}x.
\end{aligned}
\tag{5.4}
$$

Consider now the virial term $J(t)$.

Lemma 5.3.

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}J(t) \leq \int \varepsilon_1\psi_1\,\mathrm{d}x - \frac{b}{\lambda}\|\nabla_{a,\lambda}\varepsilon_0\|_{L^2}^2 + \frac{\lambda_t}{\lambda}\int [f(\varphi_0+\varepsilon_0) - f(\varphi_0)]\Lambda_0\varepsilon_0\,\mathrm{d}x \\
+ \frac{c}{t}\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}^2 + C_1 t^\gamma \cdot \|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}.
\end{aligned}
\tag{5.5}
$$

Notice the cancellation of $\int \varepsilon_1\psi_1\,\mathrm{d}x$ in (5.4) and (5.5). This is important because the bound on $\|\psi_1\|$ given by Proposition 3.14 and Proposition 4.7 is only $\frac{1}{t}\|\varepsilon\|$, which is borderline but not sufficient to close the bootstrap. Moreover, $\Lambda_0 - \Lambda = \mathrm{Id}$, so J eliminates the unbounded part of the operator Λ acting on ε_0.

Proof of Lemma 5.3. We compute

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}J(t) &= b_t\int \varepsilon_1\cdot\Big(\frac{1}{\lambda}\cdot\frac{1}{2}(\Delta a)_\lambda + (\nabla a)_\lambda\cdot\nabla\Big)\varepsilon_0\,\mathrm{d}x \\
&\quad - \frac{b\lambda_t}{\lambda}\int \varepsilon_1\cdot\Big(\frac{1}{\lambda}\cdot\frac{1}{2}(\Lambda_{3/2}\Delta a)_\lambda + (\Lambda_{5/2}\nabla a)_\lambda\cdot\nabla\Big)\varepsilon_0\,\mathrm{d}x \\
&\quad + b\int \varepsilon_{1t}\cdot\Big(\frac{1}{\lambda}\cdot\frac{1}{2}(\Delta a)_\lambda + (\nabla a)_\lambda\cdot\nabla\Big)\varepsilon_0\,\mathrm{d}x \\
&\quad + b\int \varepsilon_1\cdot\Big(\frac{1}{\lambda}\cdot\frac{1}{2}(\Delta a)_\lambda + (\nabla a)_\lambda\cdot\nabla\Big)\varepsilon_{0t}\,\mathrm{d}x.
\end{aligned}
$$

Consider the first two lines. From Lemma 5.1 and Hardy inequality it follows that

$$
\frac{1}{\lambda}\cdot\frac{1}{2}(\Delta a)_\lambda + (\nabla a)_\lambda\cdot\nabla
\tag{5.6}
$$

and

$$
\frac{1}{\lambda}\cdot\frac{1}{2}(\Lambda_{3/2}\Delta a)_\lambda - (\Lambda_{5/2}\nabla a)_\lambda\cdot\nabla
$$

are uniformly bounded as operators $\dot H^1 \to L^2$ (the bound depends on R). Moreover, it is clear that $|b_t| + \big|\frac{b\lambda_t}{\lambda}\big| \ll t^{-1}$. Hence, the first two lines are negligible.

Using again (5.1) we get

$$
\begin{aligned}
&b\int \varepsilon_{1t}\cdot\big(\frac{1}{\lambda}\cdot\frac{1}{2}(\Delta a)_\lambda+(\nabla a)_\lambda\cdot\nabla\big)\varepsilon_0\,\mathrm{d}x\\
&+b\int \varepsilon_1\cdot\big(\frac{1}{\lambda}\cdot\frac{1}{2}(\Delta a)_\lambda+(\nabla a)_\lambda\cdot\nabla\big)\varepsilon_{0t}\,\mathrm{d}x\\
&=b\int(\Delta\varepsilon_0+f(\varphi_0+\varepsilon_0)-f(\varphi_0))\cdot\big(\frac{1}{\lambda}\cdot\frac{1}{2}(\Delta a)_\lambda+(\nabla a)_\lambda\cdot\nabla\big)\varepsilon_0\,\mathrm{d}x\\
&+b\int \varepsilon_1\cdot\big(\frac{1}{\lambda}\cdot\frac{1}{2}(\Delta a)_\lambda+(\nabla a)_\lambda\cdot\nabla\big)\varepsilon_1\,\mathrm{d}x\\
&-b\int \psi_1\cdot\big(\frac{1}{\lambda}\cdot\frac{1}{2}(\Delta a)_\lambda+(\nabla a)_\lambda\cdot\nabla\big)\varepsilon_0\,\mathrm{d}x\\
&-b\int \varepsilon_1\cdot\big(\frac{1}{\lambda}\cdot\frac{1}{2}(\Delta a)_\lambda+(\nabla a)_\lambda\cdot\nabla\big)\psi_0\,\mathrm{d}x.
\end{aligned}
\tag{5.7}
$$

Proposition 3.14 and Proposition 4.7 imply that $\|\psi_1\|_{L^2}\lesssim\frac{1}{t}\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}+t^\gamma$. Using once again uniform boundedness of the operator (5.6), we obtain that the first term of the last line is negligible. Consider now the second term. We will show that

$$
\big|b\int \varepsilon_1\cdot\big(\frac{1}{\lambda}\cdot\frac{1}{2}(\Delta a)_\lambda+(\nabla a)_\lambda\cdot\nabla\big)\psi_0\,\mathrm{d}x+\int\varepsilon_1\psi_1\,\mathrm{d}x\big|\le\frac{c}{t}\|\boldsymbol{\varepsilon}\|^2_{\dot H^1\times L^2}+C_1t^\gamma\cdot\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}. \tag{5.8}
$$

It follows from Proposition 3.14 and Proposition 4.7 that in (5.8) ψ_0 can be replaced by $-(\lambda_t-b)\frac{1}{\lambda}(\Lambda W)_\lambda$ and ψ_1 by $(\lambda_t-b)\frac{b}{\lambda}(\Lambda_0\Lambda W)_{\underline{\lambda}}$. Hence, using (3.5), it suffices to prove that $\|\Lambda_0\Lambda W-\big[\frac12\Delta a+\nabla a\cdot\nabla\big]\Lambda W\|_{L^2}$ is arbitrarily small when R is large enough. But this is clear, since $\big[\frac12\Delta a+\nabla a\cdot\nabla\big]\Lambda W(r)=\Lambda_0\Lambda W(r)$ for $r\le R$ and $|\big[\frac12\Delta a+\nabla a\cdot\nabla\big]\Lambda W(r)|\lesssim r^{-3}$ for all r, with a constant independent of R.

The second line of (5.7) is 0 by integration by parts and we are left with the first line. The term with $\Delta\varepsilon_0$ is computed via a classical Pohozaev identity:

$$
\int\big[\frac{1}{\lambda}\cdot\frac{1}{2}(\Delta a)_\lambda+(\nabla a)_\lambda\cdot\nabla\big]\varepsilon_0\Delta\varepsilon_0\,\mathrm{d}x=-\frac{1}{\lambda}\|\nabla_{a,\lambda}\varepsilon_0\|^2_{L^2}+\frac{1}{4\lambda^3}\int(\Delta^2a)_\lambda\varepsilon_0^2\,\mathrm{d}x. \tag{5.9}
$$

By Lemma 5.1, the last term is finite and ≤ 0.

The nonlinear part is calculated in the following lemma.

Lemma 5.4.

$$
\begin{aligned}
\Big|b\int&\big[\frac{1}{\lambda}\cdot\frac{1}{2}(\Delta a)_\lambda+(\nabla a)_\lambda\cdot\nabla\big]\varepsilon_0\cdot[f(\varphi_0+\varepsilon_0)-f(\varphi_0)]\,\mathrm{d}x\\
&-\frac{\lambda_t}{\lambda}\int[f(\varphi_0+\varepsilon_0)-f(\varphi_0)]\Lambda_0\varepsilon_0\,\mathrm{d}x\Big|\le\frac{c}{t}\|\boldsymbol{\varepsilon}\|^2_{\dot H^1\times L^2}.
\end{aligned}
$$

We will admit for a moment that this is true and recapitulate in order to finish the proof of Lemma 5.3. Identity (5.9) implies that the term with $\Delta\varepsilon_0$ in the first line of (5.7) is *smaller* than $-\frac{b}{\lambda}\|\nabla_{a,\lambda}\varepsilon_0\|^2_{L^2}$. Lemma 5.4 implies that the difference between the other term of the first line of (5.7) and $\frac{\lambda_t}{\lambda}\int[f(\varphi_0+\varepsilon_0)-f(\varphi_0)]\Lambda_0\varepsilon_0\,\mathrm{d}x$ is negligible. The second line of (5.7) is 0, and the difference between the last line and $\int\varepsilon_1\psi_1\,\mathrm{d}x$ is negligible, as follows from the computation above. This proves (5.5). □

In order to prove Lemma 5.4, we need two auxiliary facts:

Lemma 5.5.

$$|f(k+l) - f(k) - f'(k)l - \frac{1}{2}f''(k)l^2| \lesssim |f(l)|,$$
$$|F(k+l) - F(k) - f(k)l - \frac{1}{2}f'(k)l^2| \lesssim |F(l)| + |f''(k)|l^3.$$

Proof. For $|l| \le \frac{1}{2}|k|$ this follows from the Taylor expansion and for $|l| \ge \frac{1}{2}|k|$ this is obvious by the triangle inequality. □

Lemma 5.6. *There exists a constant C_2 independent of R such that for small t,*

$$\||x| \cdot |\nabla \phi_0|\|_{L^{10/3}} \le C_2, \tag{5.10}$$
$$\||\lambda(\nabla a)_\lambda| \cdot |\nabla \phi_0|\|_{L^{10/3}} \le C_2. \tag{5.11}$$

Moreover,

$$\|\big(x - \lambda(\nabla a)_\lambda\big) \cdot \nabla \varphi_0\|_{L^{10/3}} \le c \tag{5.12}$$

if R is large enough and ρ small enough.

Proof. Recall that $\varphi_0(t) = W_{\lambda(t)} + P_0(t) + u^*(t)$, and we can estimate the three terms separately. The third one gives $\||x|\cdot|\nabla u^*|\|_{L^{10/3}}$, which is bounded by Proposition B.2 and the fact that u^* has compact support. It is easy to check that $\||x|\cdot|\nabla(W_\lambda)|\|_{L^{10/3}} = \||x|\cdot|\nabla W|\|_{L^{10/3}}$, which gives the boundedness of the first term. Finally, we compute

$$\nabla\big[\chi(\frac{x}{t})\lambda^{3/2}A_\lambda(x)\big] = \nabla\big[\chi(\frac{x}{t})A(\frac{x}{\lambda})\big] = \frac{1}{t}(\nabla\chi)(\frac{x}{t})A(\frac{x}{\lambda}) + \frac{1}{\lambda}\chi(\frac{x}{t})\nabla A(\frac{x}{\lambda}),$$

and it is sufficient to use the inequalities $|A(x/\lambda)| \lesssim \lambda/|x|$ and $|\nabla A(x/\lambda)| \lesssim \lambda^2/|x|^2$. The second term of P_0 is bounded in the same way. Notice that we obtain in fact that $\||x| \cdot |\nabla P_0(t)|\|_{L^{10/3}}$ is small when t is small.

Clearly $|\lambda(\nabla a)_\lambda| \lesssim |x|$ uniformly in R, so (5.11) follows from (5.10).

The proof of (5.12) is similar. The terms $\||x| \cdot |\nabla u^*|\|_{L^{10/3}}$ and $|\lambda(\nabla a)_\lambda| \cdot |\nabla u^*|\|_{L^{10/3}}$ are small when ρ is small. By rescaling we get

$$\|\big(x - \lambda(\nabla a)_\lambda\big) \cdot |\nabla W_\lambda|\|_{L^{10/3}} = \|(x - \nabla a) \cdot |\nabla W|\|_{L^{10/3}}.$$

By definition $\nabla a = x$ for $|x| \le R$, so

$$\|(x - \nabla a) \cdot |\nabla W|\|_{L^{10/3}} \lesssim \||x| \cdot |\nabla W|\|_{L^{10/3}(|x|\ge R)} \to 0 \qquad \text{when } R \to +\infty.$$

Smallness of $\||\big(x - \lambda(\nabla a)_\lambda\big)| \cdot |\nabla P_0(t)|\|_{L^{10/3}}$ for small t follows from smallness of $\||x| \cdot |\nabla P_0(t)|\|_{L^{10/3}}$. □

Proof of Lemma 5.4. First, as for the linear terms, using integration by parts we transform the integral so that the unbounded operator Λ_0 (and its approximation $\frac{1}{2}\Delta a + \nabla a \cdot \nabla$) no longer acts on ε_0:

$$\begin{aligned}\int \frac{1}{\lambda}x\cdot\nabla\varepsilon_0 f(\varphi_0+\varepsilon_0)\,\mathrm{d}x &= \int \frac{1}{\lambda}x\cdot\nabla(\varphi_0+\varepsilon_0)f(\varphi_0+\varepsilon_0)\,\mathrm{d}x - \int\frac{1}{\lambda}x\cdot\nabla\varphi_0 f(\varphi_0+\varepsilon_0)\,\mathrm{d}x\\ &= -5\int\frac{1}{\lambda}F(\varphi_0+\varepsilon_0)\,\mathrm{d}x - \int\frac{1}{\lambda}x\cdot\nabla\varphi_0 f(\varphi_0+\varepsilon_0)\,\mathrm{d}x\end{aligned} \tag{5.13}$$

and analogously

$$\int (\nabla a)_\lambda \cdot \nabla\varepsilon_0 f(\varphi_0+\varepsilon_0)\,\mathrm{d}x = -\int \frac{1}{\lambda}(\Delta a)_\lambda F(\varphi_0+\varepsilon_0)\,\mathrm{d}x - \int (\nabla a)_\lambda \cdot \nabla\varphi_0 f(\varphi_0+\varepsilon_0)\,\mathrm{d}x.$$

Using Lemma 5.5 we see that

$$\int \left|F(\varphi_0+\varepsilon_0) - \big(F(\varphi_0) + f(\varphi_0)\varepsilon_0 + \frac{1}{2}f'(\varphi_0)\varepsilon_0^2\big)\right|\mathrm{d}x \lesssim \|\boldsymbol{\varepsilon}\|^3_{\dot H^1\times L^2} \le f(\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2})$$

Similarly, from Lemma 5.5 and Lemma 5.6 we get

$$\int \left|x\cdot\nabla\varphi_0 f(\varphi_0+\varepsilon_0) - x\cdot\nabla\varphi_0\big(f(\varphi_0) + f'(\varphi_0)\varepsilon_0 + \frac{1}{2}f''(\varphi_0)\varepsilon_0^2\big)\right|\mathrm{d}x \lesssim f(\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}).$$

Notice that $\frac{\lambda_t}{\lambda}f(\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}) \ll \frac{1}{t}\|\boldsymbol{\varepsilon}\|^2_{\dot H^1\times L^2}$, so the above two inequalities together with (5.13) imply that

$$\begin{aligned}\frac{\lambda_t}{\lambda}\int x\cdot\nabla\varepsilon_0 f(\varphi_0+\varepsilon_0)\,\mathrm{d}x \simeq &-5\frac{\lambda_t}{\lambda}\int \big(F(\varphi_0) + f(\varphi_0)\varepsilon_0 + \frac{1}{2}f'(\varphi_0)\varepsilon_0^2\big)\,\mathrm{d}x \\ &- \frac{\lambda_t}{\lambda}\int x\cdot\nabla\varphi_0\big(f(\varphi_0) + f'(\varphi_0)\varepsilon_0 + \frac{1}{2}f''(\varphi_0)\varepsilon_0^2\big)\,\mathrm{d}x.\end{aligned} \tag{5.14}$$

Integrating by parts we find

$$\int x\cdot\nabla\varphi_0 f(\varphi_0)\,\mathrm{d}x = \int x\cdot\nabla F(\varphi_0)\,\mathrm{d}x = -5\int F(\varphi_0)\,\mathrm{d}x$$

and

$$\int x\cdot\nabla\varphi_0 f'(\varphi_0)\varepsilon_0\,\mathrm{d}x = \int x\cdot\nabla f(\varphi_0)\varepsilon_0\,\mathrm{d}x = -5\int f(\varphi_0)\varepsilon_0\,\mathrm{d}x - \int x\cdot\nabla\varepsilon_0 f(\varphi_0)\,\mathrm{d}x.$$

Thus, (5.14) simplifies to

$$\frac{\lambda_t}{\lambda}\int x\cdot\nabla\varepsilon_0\big(f(\varphi_0+\varepsilon_0) - f(\varphi_0)\big)\,\mathrm{d}x \simeq \frac{\lambda_t}{\lambda}\int\Big(-\frac{5}{2}f'(\varphi_0)\varepsilon_0^2 - \frac{1}{2}x\cdot\nabla\varphi_0 f''(\varphi_0)\varepsilon_0^2\Big)\,\mathrm{d}x. \tag{5.15}$$

Using just a pointwise estimate and Hölder we obtain

$$\frac{\lambda_t}{\lambda}\int \varepsilon_0[f(\varphi_0+\varepsilon_0) - f(\varphi_0)]\,\mathrm{d}x \simeq \frac{\lambda_t}{\lambda}\int f'(\varphi_0)\varepsilon_0^2\,\mathrm{d}x.$$

Combining with (5.15) we have

$$\begin{aligned}\frac{\lambda_t}{\lambda}\int \Lambda_0\varepsilon_0\big(f(\varphi_0+\varepsilon_0) - f(\varphi_0)\big)\,\mathrm{d}x &\simeq -\frac{\lambda_t}{2\lambda}\int x\cdot\nabla\varphi_0 f''(\varphi_0)\varepsilon_0^2\,\mathrm{d}x \\ &\simeq -\frac{b}{2\lambda}\int x\cdot\nabla\varphi_0 f''(\varphi_0)\varepsilon_0^2\,\mathrm{d}x,\end{aligned} \tag{5.16}$$

where the last almost-equality follows from the fact that $|\lambda_t - b| \lesssim \|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}$.

Analogously, we obtain

$$\begin{aligned}&b\int \big[\frac{1}{\lambda}\cdot\frac{1}{2}(\Delta a)_\lambda + (\nabla a)_\lambda\cdot\nabla\big]\varepsilon_0\big(f(\varphi_0+\varepsilon_0) - f(\varphi_0)\big)\,\mathrm{d}x \\ &\simeq -\frac{b}{2}\int (\nabla a)_\lambda\cdot\nabla\varphi_0 f''(\varphi_0)\varepsilon_0^2\,\mathrm{d}x.\end{aligned} \tag{5.17}$$

Comparing (5.16) and (5.17), we see that in order to finish the proof, we need to check that

$$\int \big|\big(x - \lambda(\nabla a)_\lambda\big) \cdot \nabla\varphi_0 f''(\varphi_0)\varepsilon_0^2\big| \,\mathrm{d}x \leq c\|\boldsymbol{\varepsilon}\|^2_{\dot H^1\times L^2}$$

when R is sufficiently large. Using Sobolev and Hölder inequalities this boils down to

$$\|(x - \lambda(\nabla a)_\lambda) \cdot \nabla\varphi_0 f''(\varphi_0)\|_{L^{5/2}} \leq c,$$

and this follows from (5.12) and boundedness of $f''(\varphi_0)$ in L^{10}. □

Proof of Proposition 5.2. From (5.4), (5.5) and the fact that $\Lambda_0 - \Lambda = \mathrm{Id}$, we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}H = \frac{\mathrm{d}}{\mathrm{d}t}I + \frac{\mathrm{d}}{\mathrm{d}t}J \leq \int \big(\Delta\varepsilon_0 + f(\varphi_0+\varepsilon_0) - f(\varphi_0)\big)\psi_0 \,\mathrm{d}x \\ -\frac{b}{\lambda}\|\nabla_{a,\lambda}\varepsilon_0\|^2_{L^2} + \frac{\lambda_t}{\lambda}\int \big(f(\varphi_0+\varepsilon_0) - f(\varphi_0)\big)\varepsilon_0\,\mathrm{d}x + \frac{c}{t}\|\boldsymbol{\varepsilon}\|^2_{\dot H^1\times L^2} + C_1 t^\gamma\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}.\end{aligned}$$

Notice that

$$\|f(\varphi_0+\varepsilon_0) - f(\varphi_0)\|_{\dot H^{-1}} \lesssim \|\varepsilon_0\|_{\dot H^1}.$$

This follows from the inequality $|f(k+l) - f(k)| \lesssim |l| + |f(l)|$ and the fact that φ_0 is bounded in $\dot H^1$. If we recall that $\frac{|\lambda_t - b|}{\lambda} \lesssim \frac{\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}}{\lambda} \ll \frac{1}{t}$, we see that in the second line we can replace λ_t by b, hence to finish the proof we only have to prove that

$$\int \big(\Delta\varepsilon_0 + f(\varphi_0+\varepsilon_0) - f(\varphi_0)\big)\psi_0\,\mathrm{d}x \leq \frac{c}{t}\|\boldsymbol{\varepsilon}\|^2_{\dot H^1\times L^2} + C_1 t^\gamma \|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}.$$

Inequalities (3.10) and (4.4) show that it is sufficient to check that

$$\Big|\int \big(\Delta\varepsilon_0 + f(\varphi_0+\varepsilon_0) - f(\varphi_0)\big)\frac{\lambda_t - b}{\lambda}(\Lambda W)_\lambda\,\mathrm{d}x\Big| \leq \frac{c}{t}\|\boldsymbol{\varepsilon}\|^2_{\dot H^1\times L^2} + C_1 t^\gamma\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2},$$

which in turn will follow from (3.5) and

$$\Big|\int \big(\Delta\varepsilon_0 + f(\varphi_0+\varepsilon_0) - f(\varphi_0)\big)(\Lambda W)_\lambda\,\mathrm{d}x\Big| \leq \frac{c\lambda}{t}\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2} + C_1\lambda t^\gamma.$$

From pointwise bounds (for example the first inequality in Lemma 5.5) one deduces

$$\|f(\varphi_0+\varepsilon_0) - f(\varphi_0) - f'(\varphi_0)\varepsilon_0\|_{\dot H^{-1}} \lesssim \|\boldsymbol{\varepsilon}\|^2_{\dot H^1} \ll \frac{\lambda}{t}\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2},$$

hence it suffices to show that

$$\Big|\int \big(\Delta\varepsilon_0 + f'(\varphi_0)\varepsilon_0\big)(\Lambda W)_\lambda\,\mathrm{d}x\Big| = \Big|\int \varepsilon_0\cdot[\Delta + f'(\varphi_0)](\Lambda W)_\lambda\,\mathrm{d}x\Big| \leq \frac{c\lambda}{t}\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2} + C_1\lambda t^\gamma.$$

Observe that $[\Delta + f'(W_\lambda)](\Lambda W)_\lambda = 0$, so we are left with proving that

$$\Big|\int \varepsilon_0\cdot(f'(\varphi_0) - f'(W_\lambda))(\Lambda W)_\lambda\,\mathrm{d}x\Big| \leq \frac{c\lambda}{t}\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2} + C_1\lambda t^\gamma.$$

By Hölder inequality it suffices to show that

$$\big\|\big(f'(\varphi_0) - f'(W_\lambda)\big)(\Lambda W)_\lambda\big\|_{L^{10/7}} \leq \frac{c\lambda}{t}. \tag{5.18}$$

The inequality $|f'(k+l) - f'(k)| \lesssim |f'(l)| + |f''(k)| \cdot |l|$ for $k = W_\lambda$ and $l = u^*(t) + P_0(t)$ reduces (5.18) to checking that

$$\|f'(u^*)(\Lambda W)_\lambda\|_{L^{10/7}} \leq \frac{c\lambda}{t}, \tag{5.19}$$

$$\|f'(P_0)(\Lambda W)_\lambda\|_{L^{10/7}} \leq \frac{c\lambda}{t}, \tag{5.20}$$

$$\|u^* \cdot f'((\Lambda W)_\lambda)\|_{L^{10/7}} \leq \frac{c\lambda}{t}, \tag{5.21}$$

$$\|P_0 \cdot f'((\Lambda W)_\lambda)\|_{L^{10/7}} \leq \frac{c\lambda}{t}. \tag{5.22}$$

Again using Hölder we get $\|P_0 \cdot f'((\Lambda W)_\lambda)\|_{L^{10/7}} \leq \|P_0\|_{L^{10/3}} \|f'((\Lambda W)_\lambda)\|_{L^{5/2}} \lesssim \|P_0\|_{\dot H^1}$. From Lemma 3.4 (or the degenerate version Lemma 4.1) we have $\|P_0\|_{\dot H^1} \lesssim t^{\gamma+1} \ll \frac{\lambda}{t}$. This proves (5.22) and (5.20) is very similar.

From Proposition B.2 we know that $\|u^*\|_{L^{10}}$ is bounded. Hence $\|u^* \cdot f'((\Lambda W)_\lambda)\|_{L^{10/7}} \lesssim \|u^*\|_{L^{10}} \cdot \|f'((\Lambda W)_\lambda)\|_{L^{5/3}} \lesssim \lambda$. This proves (5.21) and (5.19) is similar. □

6 Construction of a uniformly controlled sequence and conclusion

In this section we will analyse finite dimensional phenomena of our dynamical system – modulation equations and eigendirections of the linearized operator L. We will also define precisely the bootstrap assumptions and finish the proof of the main theorems.

It is known that the operator $L = -\Delta - f'(W)$ has a unique simple strictly negative eigenvalue $-e_0^2$ (by convention $e_0 > 0$), with a unique positive eigenfunction $\mathcal{Y}$ such that $\|\mathcal{Y}\|_{L^2} = 1$. This function $\mathcal{Y}$ is radial, smooth and decays exponentially. This follows from classical results of spectral theory and theory of elliptic equations, see [27, Proposition 5.5], where it is also shown that there exists a constant $c_1 > 0$ such that

$$g \in \dot H^1_{\mathrm{rad}}, \quad \langle g, \mathcal{Y}\rangle = \langle \nabla g, \nabla \Lambda W\rangle = 0 \quad \Rightarrow \quad \langle g, Lg\rangle \geq c_1 \|\nabla g\|_{L^2}^2. \tag{6.1}$$

We need here a slight modification of this coercivity lemma.

Lemma 6.1. *For any $c > 0$ there exist $c_L, C > 0$ such that*

$$\langle g, Lg\rangle \geq c_L \|\nabla g\|^2 - C\langle g, \mathcal{Y}\rangle^2 - c\langle g, \mathcal{Z}\rangle^2. \tag{6.2}$$

Proof. We first show that

$$g \in \dot H^1_{\mathrm{rad}}, \quad \langle g, \mathcal{Y}\rangle = \langle g, \mathcal{Z}\rangle = 0 \quad \Rightarrow \quad \langle g, Lg\rangle \geq c_2 \|\nabla g\|_{L^2}^2. \tag{6.3}$$

To prove (6.3), decompose $g = a\Lambda W + h$, $\langle h, \Delta\Lambda W\rangle = 0$. Notice that $\langle \Lambda W, \mathcal{Y}\rangle = 0$, thus $\langle h, \mathcal{Y}\rangle = 0$ and (6.1) implies

$$\langle g, Lg\rangle = \langle h + a\Lambda W, Lh\rangle = \langle h, Lh\rangle \geq c_1 \|\nabla h\|_{L^2}^2.$$

Let $\widetilde{\Lambda W}$ be the orthogonal projection of $\Delta\Lambda W$ on $\mathcal{Z}^\perp$ in $\dot H^{-1}$. We have

$$\begin{aligned}\|\nabla h\|_{L^2}^2 = \|\nabla g - a\nabla\Lambda W\|_{L^2}^2 &= \|\nabla g\|_{L^2}^2 - 2a\langle \nabla g, \nabla\Lambda W\rangle + a^2\|\nabla\Lambda W\|_{L^2}^2 \\ &= \|\nabla g\|_{L^2}^2 + 2a\langle g, \widetilde{\Lambda W}\rangle + a^2\|\nabla\Lambda W\|_{L^2}^2.\end{aligned}$$

The functions $\Delta\Lambda W$ and $\mathcal{Z}$ are not perpendicular in $\dot{H}^{-1}$, so $\|\widetilde{\Lambda W}\|_{\dot{H}^{-1}} < \|\nabla\Lambda W\|_{L^2}$, and (6.3) follows from Cauchy-Schwarz inequality.

In order to prove (6.2), we decompose

$$g = a\mathcal{Y} + b\Lambda W + \widetilde{g}, \qquad \langle\widetilde{g},\mathcal{Y}\rangle = \langle\widetilde{g},\mathcal{Z}\rangle = 0. \tag{6.4}$$

Projecting (6.4) on $\mathcal{Y}$ and $\mathcal{Z}$ we have

$$\begin{aligned} &a^2 \lesssim \langle g,\mathcal{Y}\rangle^2, \\ &b^2 \lesssim \langle g,\mathcal{Z}\rangle^2 + a^2\langle\mathcal{Z},\mathcal{Y}\rangle^2 \lesssim \langle g,\mathcal{Z}\rangle^2 + \langle g,\mathcal{Y}\rangle^2. \end{aligned} \tag{6.5}$$

From (6.3) we obtain

$$\langle\widetilde{g}, L\widetilde{f}\rangle \geq c_2\|\nabla\widetilde{g}\|_{L^2}^2,$$

thus

$$\langle g, Lg\rangle = \langle a\mathcal{Y} + b\Lambda W + \widetilde{g}, -e_0^2 a\mathcal{Y} + L\widetilde{g}\rangle = -e_0^2a^2 + \langle\widetilde{g}, L\widetilde{g}\rangle \geq c_2\|\nabla\widetilde{g}\|_{L^2}^2 - e_0^2a^2.$$

From the inequality $(x-y)^2 \geq \frac{1}{2}x^2 - y^2$ we have

$$\|\nabla\widetilde{g}\|_{L^2}^2 \geq \frac{1}{2}\|\nabla g - b\nabla\Lambda W\|_{L^2}^2 - a^2\|\nabla\mathcal{Y}\|_{L^2}^2.$$

From the inequality $(x-y)^2 \geq \frac{c}{1+c}x^2 - cy^2$ we have

$$\|\nabla g - b\nabla\Lambda W\|_{L^2}^2 \geq \frac{c}{1+c}\|\nabla g\|_{L^2}^2 - cb^2\|\nabla\Lambda W\|_{L^2}^2.$$

If we choose c small enough and put everything together using (6.5), we obtain (6.2). $\square$

From now on we will denote

$$\alpha(g) := \langle g,\mathcal{Y}\rangle, \qquad \alpha_\lambda(g) := \Big\langle g, \frac{1}{\lambda}\mathcal{Y}_{\underline{\lambda}}\Big\rangle.$$

We prove a version of the coercivity lemma with a localized gradient term.

Lemma 6.2. *Let $c > 0$. If R is large enough, then there exists a constant C such that*

$$\int_{|x|\leq R}|\nabla g|^2\,\mathrm{d}x - \int_{\mathbb{R}^5} f'(W)g^2\,\mathrm{d}x \geq -c\|\nabla g\|_{L^2}^2 - C|\alpha(g)|^2.$$

In the proof we assume that g is radial, which is justified because later we use it for $g = \varepsilon_0$. Notice however that the non-radial case follows by considering the radial rearrangement.

Proof. Define the projection $\Psi_R : \dot{H}^1 \to \dot{H}^1$ by the formula:

$$\Psi_R g(r) = \begin{cases} g(r) - g(R) & \text{if } r \leq R, \\ 0 & \text{if } r \geq R. \end{cases}$$

By (6.2) applied to $\Psi_R g$ we have

$$\begin{aligned} \Big(1+\frac{c}{2}\Big)\int_{|x|\leq R}|\nabla g|^2\,\mathrm{d}x &= \Big(1+\frac{c}{2}\Big)\int_{\mathbb{R}^5}|\nabla(\Psi_R g)|^2\,\mathrm{d}x \\ &\geq \Big(1+\frac{c}{2}\Big)\int f'(W)|\Psi_R g|^2\,\mathrm{d}x - C\langle\Psi_R g,\mathcal{Y}\rangle^2 - \frac{c}{4}\langle\Psi_R g,\mathcal{Z}\rangle^2. \end{aligned}$$

Recall that, by the Strauss Lemma [87], in dimension $N = 5$ for a radial function g we have $|g(R)| \lesssim C_0 R^{-\frac{3}{2}}\|\nabla g\|_{L^2}$ with a universal constant C_0, so we have a pointwise estimate

$$|g|^2 \leq C_0\big(1+\frac{2}{c}\big)R^{-3}\|\nabla g\|_{L^2}^2 + \big(1+\frac{c}{2}\big)|\Psi_R g|^2.$$

Now we notice that

$$\int_{|x|\leq R} f'(W)\,\mathrm{d}x \sim R,$$

so for any $\delta > 0$ the first term above gives a small contribution to the quadratic form for R large. Similarly,

$$|\langle g - \Psi_R g, \mathcal{Y}\rangle| \lesssim R^{-\frac{3}{2}}\|\nabla g\|_{L^2} + \int_{|x|\geq R}|g|\mathcal{Y}\,\mathrm{d}x \leq \big(R^{-\frac{3}{2}} + \|\mathcal{Y}\|_{L^{10/7}(|x|\geq R)}\big)\|\nabla g\|_{L^2},$$

which is small when R is large. As $\langle \Psi_R g, \mathcal{Y}\rangle^2 \leq 2\langle g, \mathcal{Y}\rangle^2 + 2\langle g - \Psi_R g, \mathcal{Y}\rangle^2$, the proof is finished. □

We are ready to state coercivity properties of the functional H from the previous section.

Proposition 6.3. *Under the assumptions of Proposition 5.2, there exist $T_0, c_H, \alpha_0, C_2 > 0$ such that for $t \in (0, T_0] \cap (T_-, T_+)$ there holds*

$$|\alpha_\lambda(\varepsilon_0)| \leq \alpha_0\|\varepsilon_0\|_{\dot H^1} \quad \Rightarrow \quad H(t) \geq c_H\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}^2.$$

If $[T_1, T_2] \subset (0, T_0] \cap (T_-, T_+)$ and $|\alpha_\lambda(\varepsilon_0)| \leq \frac{1}{e_0}t^{\gamma+1}$ for all $t \in [T_1, T_2]$, then

$$H(T_2) \leq H(T_1) + \frac{c_H}{10}\int_{T_1}^{T_2}\frac{1}{t}\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}^2\,\mathrm{d}t + C_2 t^{2\gamma+2}. \tag{6.6}$$

The constants $\frac{c_H}{10}$ and $\frac{1}{e_0}$ have no special signification, but this formulation will be convenient later.

Proof. Let

$$I_{\text{LIN}}(t) := \int \frac{1}{2}|\varepsilon_1|^2 + \frac{1}{2}|\nabla\varepsilon_0|^2 - \frac{1}{2}f'(W_\lambda)\varepsilon_0^2\,\mathrm{d}x.$$

Recall that $\langle \varepsilon_0, \mathcal{Z}\rangle = 0$. Lemma 6.1 implies (after rescaling) that if we take α_0 small enough, then there exists a constant $c > 0$ such that $I_{\text{LIN}}(t) \geq c\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}^2$.

We can assume that $\|\boldsymbol{u}^*\|_{X^1\times H^1}$ is as small as we like, so by pointwise estimates we get $|I(t) - I_{\text{LIN}}(t)| \leq \frac{1}{3}c\|\boldsymbol{\varepsilon}_0\|_{\dot H^1\times L^2}^2$. Moreover, it is clear from the definition of J that for small t we have $|J(t)| \leq \frac{1}{3}c\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}^2$. This proves the result with $c_H = \frac{1}{3}c$.

In order to prove (6.6), notice first that, by pointwise estimates and smallness of $\|\varphi_0 - W_\lambda\|_{\dot H^1}$, in (5.2) we can replace $\int \big(f(\varphi_0+\varepsilon_0) - f(\varphi_0)\big)\varepsilon_0\,\mathrm{d}x$ by $\int f'(W_\lambda)\varepsilon_0^2\,\mathrm{d}x$. Convexity of a (see Lemma 5.1) implies that

$$\|\nabla_{a,\lambda}\varepsilon\|_{L^2}^2 \geq \int_{|x|\leq R\lambda}|\nabla\varepsilon_0|^2\,\mathrm{d}x,$$

so from (5.2) and Lemma 6.2 (after rescaling) we obtain

$$\begin{aligned} H(T_2) &\leq H(T_1) + \int_{T_1}^{T_2}\frac{c_H}{20t}\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}^2 + C_1 t^\gamma\cdot\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2} + C|\alpha(\varepsilon_0)|^2\,\mathrm{d}t \\ &\leq H(T_1) + \frac{c_H}{10}\int_{T_1}^{T_2}\frac{1}{t}\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2}^2\,\mathrm{d}t + C_2 t^{2\gamma+2}, \end{aligned}$$

where C_2 is a constant. □

In order to close the bootstrap, it is necessary to control the stable and unstable directions. More precisely, it is necessary to eliminate the unstable mode.

Define

$$\alpha_\lambda^-(\boldsymbol{\varepsilon}) := \int \mathcal{Y}_{\underline{\lambda}}\varepsilon_1 - \frac{e_0}{\lambda}\mathcal{Y}_{\underline{\lambda}}\varepsilon_0 \,\mathrm{d}x$$

and

$$\alpha_\lambda^+ := \int \mathcal{Y}_{\underline{\lambda}}\varepsilon_1 + \frac{e_0}{\lambda}\mathcal{Y}_{\underline{\lambda}}\varepsilon_0 \,\mathrm{d}x.$$

Notice that $-\frac{e_0^2}{\lambda^2}$ is the unique strictly negative eigenvalue of L_λ.

We will define an auxiliary function $l(t)$ which measures the distance of the modulation parameters from the approximate trajectory (2.7) or (2.8). This function has a slightly different form in the non-degenerate and degenerate cases. In the non-degenerate case we define

$$l(t) = \frac{1}{2}\Big(\frac{b}{t^3} + \frac{2\lambda}{t^4} - \frac{\kappa^2 u^*(0,0)^2}{24}\Big)^2 + \frac{1}{2}\Big(\frac{b}{t^3} - \frac{3\lambda}{t^4} - \frac{\kappa^2 u^*(0,0)^2}{144}\Big)^2,$$

and in the degenerate case

$$l(t) = \frac{1}{2}\big(\frac{b}{t^\nu} + \widetilde{\nu}\frac{\lambda}{t^{\nu+1}} - (\nu + \widetilde{\nu} + 1)\big)^2 + \frac{1}{2}\big(\frac{b}{t^\nu} - (\widetilde{\nu}+1)\frac{\lambda}{t^{\nu+1}} - (\nu - \widetilde{\nu})\big)^2,$$

where $\widetilde{\nu} := -\frac{1}{2} + \frac{1}{2}\sqrt{\nu^2 + (\nu+1)^2}$.

We will write $\alpha^+(t)$ and $\alpha^-(t)$ instead of $\alpha_\lambda^+(\boldsymbol{\varepsilon})$ and $\alpha_\lambda^-(\boldsymbol{\varepsilon})$. In the next few propositions we describe the evolution of $\mathrm{Mod}(t) := (\lambda(t), b(t), \alpha^-(t), \alpha^+(t))$ in the "modulation cylinder" defined as:

$$\mathscr{C}(t) := \big\{(\lambda, b, \alpha^-, \alpha^+) : \ l(t) \le t^{\gamma+1-\nu} \text{ and } -t^{\gamma+1} \le \alpha^-, \alpha^+ \le t^{\gamma+1}\big\}.$$

In the non-degenerate case we denote

$$\lambda_{\mathrm{app}}(t) := \frac{\kappa^2 u^*(0,0)^2}{144}t^4, \qquad b_{\mathrm{app}}(t) := \frac{\kappa^2 u^*(0,0)^2}{36}t^3,$$

and in the degenerate case

$$\lambda_{\mathrm{app}}(t) := t^{\nu+1}, \qquad b_{\mathrm{app}}(t) := (\nu+1)t^\nu.$$

Solving a 2×2 linear system we check easily that

$$l(t) \le t^{\gamma+1-\nu} \Rightarrow \big|\frac{\lambda}{\lambda_{\mathrm{app}}} - 1\big| \lesssim t^{\frac{1}{2}(\gamma+1-\nu)}\ ,\ \big|\frac{b}{b_{\mathrm{app}}} - 1\big| \lesssim t^{\frac{1}{2}(\gamma+1-\nu)}, \tag{6.7}$$

with constants which depend only on ν.

We have $\alpha_\lambda(\varepsilon_0) = \frac{1}{2e_0}(\alpha^+ - \alpha^-)$, so

$$\mathrm{Mod}(t) \in \mathscr{C}(t) \quad \Rightarrow \quad |\alpha_\lambda(\varepsilon_0)| \le \frac{1}{e_0}t^{\gamma+1}. \tag{6.8}$$

Remark 6.4. The formula for l is found by linearizing the parameter equations near $(\lambda_{\mathrm{app}}, b_{\mathrm{app}})$ and diagonalizing the resulting system.

We can finally state a result on uniform in time energy bounds.

Proposition 6.5. *Let $C_4 > 0$ be a fixed constant. There exist $C_0 > 0$ and $T_0 > 0$ having the following property. Let $0 < T_1 < T_0$. Suppose that*

$$\|\boldsymbol{\varepsilon}(T_1)\|_{\dot H^1\times L^2} \leq C_4 T_1^{\gamma+1}, \qquad (6.9)$$
$$\mathrm{Mod}(T_1) \in \mathrm{Int}(\mathscr{C}(T_1)).$$

Then, either there exists a time t, $T_1 \leq t \leq T_0$, such that $\mathrm{Mod}(t) \in \partial\mathscr{C}(t)$*, or the solution exists on $[T_1, T_0]$ and for all $t \in [T_1, T_0]$ there holds*

$$\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2} \leq C_0 t^{\gamma+1}. \qquad (6.10)$$

Proof. Let T_0 be the time provided by Proposition 6.3. Let T_+ be the maximal time of existence of the solution and let $T_2 := \min(T_0, T_+)$. Suppose that $\mathrm{Mod}(t) \notin \partial\mathscr{C}(t)$ for $T_1 \leq t \leq T_2$. By continuity of $\mathrm{Mod}(t)$ this means that $\mathrm{Mod}(t) \in \mathrm{Int}(\mathscr{C}(t))$ for $T_1 \leq t \leq T_2$. We will show first that if C_0 is large enough, then (6.10) holds for $t \in [T_1, T_2]$. Argue by contradiction, assuming that there exists $T_3 < T_2$ such that $\|\boldsymbol{\varepsilon}(T_3)\|_{\dot H^1\times L^2} = C_0 T_3^{\gamma+1}$. At $t = T_3$ (6.8) gives $|\alpha_\lambda(\varepsilon_0)| \leq \frac{1}{e_0} t^{\gamma+1}$. In particular, if C_0 is large, we will have $|\alpha_\lambda(\varepsilon_0)| \leq \alpha_0\|\varepsilon_0\|_{\dot H^1}$, so by Proposition 6.3 we obtain

$$H(T_3) \geq c_H C_0^2 T_3^{2\gamma+2}. \qquad (6.11)$$

On the other hand, for $t \in [T_1, T_3]$ we have $\|\boldsymbol{\varepsilon}\|^2_{\dot H^1\times L^2} \leq C_0^2 t^{2\gamma+2}$ and $|\alpha_\lambda(\varepsilon_0)| \leq \frac{1}{e_0} t^{\gamma+1}$, so from (6.6) we deduce that

$$H(T_3) \leq H(T_1) + \frac{c_H C_0^2}{10(2\gamma+2)} T_3^{2\gamma+2} + C_2 T_3^{2\gamma+2}.$$

Notice that $H(T_1) \leq \|\boldsymbol{\varepsilon}(T_1)\|^2_{\dot H^1\times L^2} \leq C_4^2 T_1^{2\gamma+2} \leq C_4^2 T_3^{2\gamma+2}$. Returning to (6.11) we deduce

$$c_H C_0^2 \leq C_4^2 + \frac{c_H C_0^2}{10(2\gamma+2)} + C_2,$$

which is impossible if C_0 is large enough.

Hence, $T_3 = T_2$. To prove that $T_2 = T_0$, notice that by the Cauchy theory in the critical space there exists $\delta > 0$ such that

$$\|(u_0 - W, u_1)\|_{\dot H^1\times L^2} \leq \delta \Rightarrow \begin{array}{l}\text{the solution } \boldsymbol{u}(t) \text{ with } \boldsymbol{u}(0) = (u_0, u_1)\\ \text{exists at least for } t \in (-1, 1).\end{array}$$

After rescaling we obtain

$$\|(u_0 - W_\lambda, u_1)\|_{\dot H^1\times L^2} \leq \delta \Rightarrow \begin{array}{l}\text{the solution } \boldsymbol{u}(t) \text{ with } \boldsymbol{u}(0) = (u_0, u_1)\\ \text{exists at least for } t \in (-\lambda, \lambda).\end{array} \qquad (6.12)$$

If $\|\boldsymbol{u}^*\|_{\dot H^1\times L^2}$ is sufficiently small and T_0 is chosen sufficiently small, (3.2) and (3.4) show that our solution verifies the sufficient condition in (6.12) for any $t < T_2$ with $\lambda = \lambda(t)$. Taking t close to T_2 we obtain that the solution cannot blow up at T_2, hence $T_2 = T_0$. □

The crucial element of the preceding result is that the constant C_0 is independent of T_1. From now, C_0 has a fixed value given by Proposition 6.5, and the constants which appear later are allowed to depend on C_0. In particular, when we use the notation $\lesssim$ or O, the constant may depend on C_0.

We examine now the evolution of the eigenvectors α^- and α^+.

Lemma 6.6. *If $\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2} \lesssim t^{\gamma+1}$, $\lambda \sim t^{\nu+1}$ and $b \sim t^\nu$, then*

$$\Big|\frac{\mathrm{d}}{\mathrm{d}t}\alpha_\lambda^+ - \frac{e_0}{\lambda}\alpha_\lambda^+\Big| \lesssim t^\gamma, \tag{6.13}$$

$$\Big|\frac{\mathrm{d}}{\mathrm{d}t}\alpha_\lambda^- + \frac{e_0}{\lambda}\alpha_\lambda^-\Big| \lesssim t^\gamma. \tag{6.14}$$

Proof. We will do the computation for (6.13), because the one for (6.14) is exactly the same.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\alpha_\lambda^+ = &\int \big(\mathcal{Y}_{\underline{\lambda}}\cdot(-L_\lambda\varepsilon_0) + \frac{e_0}{\lambda}\mathcal{Y}_{\underline{\lambda}}\varepsilon_1\big)\,\mathrm{d}x\\ &+\int \frac{-\lambda_t}{\lambda}\big((\Lambda_0\mathcal{Y})_{\underline{\lambda}}\varepsilon_1 + \frac{e_0}{\lambda}(\Lambda_{-1}\mathcal{Y})_{\underline{\lambda}}\varepsilon_0\big)\,\mathrm{d}x\\ &+\int \mathcal{Y}_{\underline{\lambda}}\cdot\big(f(\varphi_0+\varepsilon_0) - f(\varphi_0) - f'(\varphi_0)\varepsilon_0\big)\,\mathrm{d}x\\ &+\int \mathcal{Y}_{\underline{\lambda}}\cdot\big(f'(\varphi_0) - f'(W_\lambda)\big)\varepsilon_0\,\mathrm{d}x\\ &+\int \big(\mathcal{Y}_{\underline{\lambda}}\cdot(-\psi_1) + \frac{e_0}{\lambda}\mathcal{Y}_{\underline{\lambda}}\cdot(-\psi_0)\big)\,\mathrm{d}x.\end{aligned}$$

The first line is $\frac{e_0}{\lambda}\alpha_\lambda^+$ and it suffices to estimate the remaining ones. For the last line we use Proposition 3.14 and L^2-orthogonality of ΛW and $\mathcal{Y}$. Using $\lambda_t \sim t^\nu$, $\lambda \sim t^{\nu+1}$ and $\|\boldsymbol{\varepsilon}\|_{\dot H^1\times L^2} \leq Ct^{\gamma+1}$ the second line is seen to be bounded by Ct^γ. The proof of (5.18) shows that $\|\mathcal{Y}_{\underline{\lambda}}\cdot\big(f'(\varphi_0) - f'(W_\lambda)\big)\|_{L^{10/7}} \leq \frac{c}{t}$, so using $\|\varepsilon_0\|_{L^{10/3}} \lesssim t^{\gamma+1}$ we obtain the required bound for the fourth line. Finally, $\|f(\varphi_0+\varepsilon_0) - f(\varphi_0) - f'(\varphi_0)\varepsilon_0\|_{\dot H^{-1}} \lesssim C^2t^{2\gamma+2}$ and $\|\mathcal{Y}_{\underline{\lambda}}\|_{\dot H^1} \lesssim \frac{1}{\lambda} \sim t^{-\nu-1}$, so by Cauchy-Schwarz the third line is bounded by $C^2t^{2\gamma-\nu+1} \ll t^\gamma$. □

We know from Proposition 6.5 that if we start at $t = T_1$ with $\boldsymbol{\varepsilon}$ small enough, then $\boldsymbol{\varepsilon}$ is controlled in $\dot H^1 \times L^2$ unless Mod leaves the cylinder $\mathscr{C}$. It turns out that it can happen only because of α^+. The other parameters are trapped in the cylinder for small times:

Lemma 6.7. *Under the assumptions of Propositon 6.5, suppose that* $\mathrm{Mod}(t)$ *leaves* $\mathrm{Int}(\mathscr{C}(t))$ *before $t = T_0$. If $T_2 \leq T_0$ is the first time for which* $\mathrm{Mod}(T_2) \in \partial\mathscr{C}(T_2)$*, then $|\alpha^+(T_2)| = T_2^{\gamma+1}$. In addition, suppose that at time T_3, $T_1 \leq T_3 < T_2$, we have $\alpha^+(T_3) > \frac{1}{2}T_3^{\gamma+1}$. Then $\alpha^+(T_2) = T_2^{\gamma+1}$. Analogously, if $\alpha^+(T_3) < -\frac{1}{2}T_3^{\gamma+1}$, then $\alpha^+(T_2) = -T_2^{\gamma+1}$.*

Proof. Suppose, for the sake of contradiction, that for example $l(T_2) = T_2^{\gamma+1-\nu}$. In particular this implies $\frac{\mathrm{d}}{\mathrm{d}t}l(t_1) \geq 0$, and we will show that it is impossible.

We start with the degenerate case. Using (6.7) and $\sqrt{x} = \frac{1+x}{2} + O(|1-x|^2)$ we obtain

$$\sqrt{\lambda} = \frac{1}{2}\big(t^{\frac{1}{2}(\nu+1)} + \lambda t^{-\frac{1}{2}(\nu+1)}\big) + O(t^{\gamma+\frac{3}{2}-\frac{1}{2}\nu}).$$

Recall that $b_t = (\nu+1)\nu t^{\frac{1}{2}(\nu-3)}\sqrt{\lambda}$, so we get

$$\frac{b_t}{t^{\nu-1}} = \frac{(\nu+1)\nu}{2}\big(1 + \frac{\lambda}{t^{\nu+1}}\big) + O(t^{\gamma+1-\nu}). \tag{6.15}$$

From Lemma 3.7 and (6.10) we have

$$\frac{\lambda_t}{t^\nu} = \frac{b}{t^\nu} + O(t^{\gamma+1-\nu}). \tag{6.16}$$

Using (6.15) and (6.16) we can compute $\frac{\mathrm{d}}{\mathrm{d}t}l(t)$:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}l(t) &= \big(\frac{b}{t^\nu} + \widetilde{\nu}\frac{\lambda}{t^{\nu+1}} - (\nu+\widetilde{\nu}+1)\big)\big(\frac{b_t}{t^\nu} + \widetilde{\nu}\frac{\lambda_t}{t^{\nu+1}} - \nu\frac{b}{t^{\nu+1}} - \widetilde{\nu}(\nu+1)\frac{\lambda}{t^{\nu+2}}\big)\\
&+ \big(\frac{b}{t^\nu} - (\widetilde{\nu}+1)\frac{\lambda}{t^{1+\nu}} - (\nu-\widetilde{\nu})\big)\big(\frac{b_t}{t^\nu} - (\widetilde{\nu}+1)\frac{\lambda_t}{t^{\nu+1}} - \nu\frac{b}{t^{\nu+1}} - (\widetilde{\nu}+1)(\nu+1)\frac{\lambda}{t^{\nu+2}}\big)\\
&= \frac{1}{t}\big(\frac{b}{t^\nu} + \widetilde{\nu}\frac{\lambda}{t^{\nu+1}} - (\nu+\widetilde{\nu}+1)\big)\big(\tfrac{(\nu+1)\nu}{2}\big(1+\tfrac{\lambda}{t^{\nu+1}}\big) + \widetilde{\nu}\tfrac{b}{t^\nu} - \nu\tfrac{b}{t^\nu} - \widetilde{\nu}(\nu+1)\tfrac{\lambda}{t^{\nu+1}}\big)\\
&+ \big(\frac{b}{t^\nu} - (\widetilde{\nu}+1)\frac{\lambda}{t^{1+\nu}} - (\nu-\widetilde{\nu})\big)\big(\tfrac{(\nu+1)\nu}{2}\big(1+\tfrac{\lambda}{t^{\nu+1}}\big) - (\widetilde{\nu}+1)\tfrac{b}{t^\nu} - \nu\tfrac{b}{t^\nu} - (\widetilde{\nu}+1)(\nu+1)\tfrac{\lambda}{t^{\nu+1}}\big)\\
&+ \frac{1}{t}O(\sqrt{l(t)}t^{\gamma+1-\nu}).\end{aligned}$$

If we use the definition of $\widetilde{\nu}$, this simplifies to

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}l(t) = &- (\nu-\widetilde{\nu})\big(\frac{b}{t^\nu} + \widetilde{\nu}\frac{\lambda}{t^{\nu+1}} - (\nu+\widetilde{\nu}+1)\big)^2 - (\nu+\widetilde{\nu}+1)\big(\frac{b}{t^\nu} - (\widetilde{\nu}+1)\frac{\lambda}{t^{1+\nu}} - (\nu-\widetilde{\nu})\big)^2\\
&+ \frac{1}{t}O(\sqrt{l(t)}t^{\gamma+1-\nu}) = \frac{1}{t}\big(-(\nu-\widetilde{\nu})l(t) + O(t^{\frac{3}{2}(\gamma+1-\nu)})\big).\end{aligned}$$

At time $t = T_2$ by assumption $l(T_2) = T_2^{\gamma+1-\nu}$, so for T_2 small enough the formula above implies $\frac{\mathrm{d}}{\mathrm{d}t}l(T_2) < 0$, which is impossible.

In the non-degenerate case the computation is similar, but we must take into account that in this case

$$b_t = \kappa u^*(t,0)\sqrt{\lambda} = \kappa u^*(0,0)\sqrt{\lambda}(1+O(t)),$$

which leads to

$$b_t = \kappa u^*(0,0)\cdot\frac{1}{2}\big(\frac{\kappa u^*(0,0)}{12}t^2 + \frac{12}{\kappa u^*(0,0)}\lambda t^{-2}\big) + O(t^3).$$

Then, the computation is the same as before:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}l(t) &= \big(\frac{b}{t^3} + \frac{2\lambda}{t^4} - \frac{\kappa^2u^*(0,0)^2}{24}\big)\big(\frac{b_t}{t^3} + \frac{2\lambda_t}{t^4} - \frac{3b}{t^4} - \frac{8\lambda}{t^5}\big)\\
&+ \big(\frac{b}{t^3} + \frac{2\lambda}{t^4} - \frac{\kappa^2u^*(0,0)^2}{24}\big)\big(\frac{b_t}{t^3} + \frac{2\lambda_t}{t^4} - \frac{3b}{t^4} - \frac{8\lambda}{t^5}\big)\\
&\leq \frac{1}{t}\big(\frac{b}{t^\nu} + \frac{2\lambda}{t^4} - \frac{\kappa^2u^*(0,0)^2}{24}\big)\big(\frac{\kappa^2u^*(0,0)^2}{24} + \frac{6\lambda}{t^4} + \frac{2b}{t^3} - \frac{3b}{t^3} - \frac{8\lambda}{t^4}\big)\\
&+ \frac{1}{t}\big(\frac{b}{t^3} - \frac{3\lambda}{t^4} - \frac{\kappa^2u^*(0,0)^2}{144}\big)\big(\frac{\kappa^2u^*(0,0)^2}{24} + \frac{6\lambda}{t^4} - \frac{3b}{t^3} - \frac{3b}{t^3} + \frac{12\lambda}{t^4}\big) + O(\sqrt{l(t)})\\
&\leq \frac{1}{t}\big(-2l(t) + O(t^{7/4})\big).\end{aligned}$$

Since $\gamma+1-\nu = \frac{3}{2} < \frac{7}{4}$, we are done.

Now suppose that $|\alpha^-(T_2)| = T_2^{\gamma+1}$. As $\frac{t^{\gamma+1}}{\lambda} \sim t^{\gamma-\nu} \gg t^\gamma$, (6.14) implies that $\frac{\mathrm{d}}{\mathrm{d}t}\alpha_\lambda^-$ and α_λ^- have opposite signs, which is impossible.

Again by contradiction, suppose that $\alpha_\lambda^+(T_3) > \frac{1}{2}T_3^{\gamma+1}$ and $\alpha_\lambda^+(T_2) = -T_2^{\gamma+1}$. By continuity, there exists the smallest $T_4 > T_3$ such that $\alpha_\lambda^+(T_4) = \frac{1}{2}T_4^{\gamma+1}$. Necessarily $\frac{\mathrm{d}}{\mathrm{d}t}\alpha_\lambda^+(T_4) \leq \frac{\gamma+1}{2}T_4^\gamma$, which is in contradiction with (6.13). □

Proposition 6.8. *There exist strictly positive constants C_0 and T_0 such that for all $T_1 \in (0,T_0)$ there exists a solution $\boldsymbol{u}$ defined on $[T_1,T_0]$ which for all $t \in [T_1,T_0]$ verifies*

$$\|(u - W_\lambda - u^*, \partial_t u + \lambda_t(\Lambda W)_{\underline{\lambda}} - \partial_t u^*)\|_{\dot H^1\times L^2} \leq C_0 t^{\gamma+1}, \tag{6.17}$$

$$\Big|\frac{\lambda}{\lambda_{\mathrm{app}}} - 1\Big| \leq C_0 t^{\frac{1}{2}(\gamma+1-\nu)}. \tag{6.18}$$

Proof. We consider the degenerate case. The proof in the non-degenerate case is similar.

Let $\lambda = \lambda_{\mathrm{app}}(T_1)$, $b = b_{\mathrm{app}}(T_1)$. For $a \in [-\frac{2}{3}T_1^{\gamma+1}, \frac{2}{3}T_1^{\gamma+1}]$, let $\boldsymbol{\varepsilon}_a(T_1) = \frac{a}{2\nu}(\mathcal{Y}_\lambda^+ - \frac{\langle \mathcal{Y}, \mathcal{Z}\rangle}{\langle \mathcal{Z}, \mathcal{Z}\rangle}(\mathcal{Z}_{\underline{\lambda}}, 0))$, and consider the corresponding evolution. Of course (6.9) is verified for a universal constant C_4. Let C_0 be the constant provided by Proposition 6.5. We will show that there exists a parameter a for which the solution exists until $t = T_0$ and satisfies (6.10). Suppose this is not the case. Let $\mathcal{A}^+ = \{a : \alpha^+(T_2) = T_2^{\gamma+1}\}$ and $\mathcal{A}^- = \{a : \alpha^+(T_2) = -T_2^{\gamma+1}\}$, where T_2 is the exit time given by Lemma 6.7. By the second part of Lemma 6.7 we know that $-\frac{2}{3}T_1^{\gamma+1} \in \mathcal{A}^-$, $\frac{2}{3}T_1^{\gamma+1} \in \mathcal{A}^+$, and that $\mathcal{A}^-$ and $\mathcal{A}^+$ are open sets. Indeed, let $a \in \mathcal{A}^+$. This means in particular that for $T_1 \leq t \leq T_2$ we have $\alpha^+(t) \geq -\frac{1}{2}t^{\gamma+1}$ and $\alpha^+(T_2) = T_2^{\gamma+1}$. By continuity of the flow, for close enough initial data we will still have $\alpha^+(t) > -t^{\gamma+1}$ for $T_1 \leq t \leq T_2$ and $\alpha^+(T_2) \geq \frac{1}{2}T_2^{\gamma+1}$. By Lemma 6.7 the corresponding solutions escape from the cylinder by positive values of α^+. Thus $\mathcal{A}^+ \cup \mathcal{A}^-$ would be a partition of $[-\frac{2}{3}T_1^{\gamma+1}, \frac{2}{3}T_1^{\gamma+1}]$ into two disjoint open sets, which is impossible.

Using (6.10), (4.2), (4.3) and (3.5) we obtain (6.17).

Estimate (6.18) follows from (6.7) and the fact that $\mathrm{Mod}(t) \in \mathscr{C}(t)$. □

Proof of Theorem 1. Let t_n be a decreasing sequence such that $t_n > 0$ and $t_n \to 0$. Let $\boldsymbol{u}_n$ be the solution given by Proposition 6.8 for $T_1 = t_n$ and let $\lambda_n : [t_n, T_0] \to (0, +\infty)$ be the corresponding modulation parameter. The sequence $\boldsymbol{u}_n(T_0)$ is bounded in $\dot H^1 \times L^2$. After extracting a subsequence, it converges weakly to some function (u_0, u_1). Let $\boldsymbol{u}(t)$ be the solution of (NLW) for the Cauchy data $\boldsymbol{u}(T_0) = (u_0, u_1)$. We will show that $\boldsymbol{u}$ satisfies (1.1).

Let $0 < T_1 < T_0$ and $T_1 \leq t \leq T_0$. Using (6.17), (6.18) and $|\lambda_t| \lesssim t^3$ we get

$$\|(u_n - W_{\lambda_{\mathrm{app}}} - u^*, \partial_t u_n - \partial_t u^*)\|_{\dot H^1 \times L^2} \leq C_0 t^{\frac{3}{4}}.$$

This shows that if T_0 is sufficiently small, then the sequence $\boldsymbol{u}_n$ satisfies the conditions of Proposition A.1 on the time interval $[T_1, T_0]$, hence

$$\boldsymbol{u}_n(T_1) \rightharpoonup \boldsymbol{u}(T_1).$$

Weak lower semi-continuity of the norm implies that at time $t = T_1$ we have

$$\|(u - W_{\lambda_{\mathrm{app}}} - u^*, \partial_t u - \partial_t u^*)\|_{\dot H^1 \times L^2} \leq C_0 T_1^{3/4}.$$

This bound holds for all T_1 such that $0 < T_1 < T_0$. In particular, the orthogonality condition:

$$\langle u - W_\lambda - u^*, \mathcal{Z}_{\underline{\lambda}}\rangle = 0. \tag{6.19}$$

defines uniquely a continuous function $\lambda(T_1) : (0, T_0) \to (0, +\infty)$. We will prove that $\lambda_n(T_1) \to \lambda(T_1)$.

Using (3.7) for the solution u_n at time T_1 and passing to a limit $n \to \infty$ we obtain that all the accumulation points of $\lambda_n(T_1)$ verify the orthogonality condition (6.19). Hence $\lambda_n(T_1) \to \lambda(T_1)$. Passing to a limit in (3.8) we get $\frac{\mathrm{d}}{\mathrm{d}t}\lambda_n(T_1) \to \frac{\mathrm{d}}{\mathrm{d}t}\lambda(T_1)$. Passing to a limit in (6.17) and (6.18) finishes the proof. □

The proof of Theorem 2 follows the same lines, so we will skip it.

A Weak continuity of the flow near a fixed path

Proposition A.1. *Let $\boldsymbol{v} : [0,1] \to \dot{H}^1 \times L^2$ be a continuous path in the energy space. There exist a constant $\delta > 0$ with the following property. Let $\boldsymbol{u}_n$ be a sequence of radial solutions of* (NLW) *defined on the interval* $[0,1]$, *such that*

$$\sup_{t\in[0,1]} \|\boldsymbol{u}_n - \boldsymbol{v}\|_{\dot{H}^1\times L^2} \leq \delta. \tag{A.1}$$

Suppose that $\boldsymbol{u}_n(0) \rightharpoonup (u_0, u_1)$ in $\dot{H}^1 \times L^2$ and let $\boldsymbol{u}$ be the solution of (NLW) *for the initial data $\boldsymbol{u}(0) = (u_0, u_1)$. Then $\boldsymbol{u}$ is defined on $[0,1]$ and for all $t \in [0,1]$ we have*

$$\boldsymbol{u}_n(t) \rightharpoonup \boldsymbol{u}(t) \qquad \text{in } \dot{H}^1 \times L^2. \tag{A.2}$$

Remark A.2. Notice that without the assumption (A.1) the result is false. More generally, existence of type II blow-up solutions in some space excludes weak continuity of the flow in this space, and existence of type II blowup solutions in our case follows from Theorems 1 and 2. One might search weaker conditions than (A.1); we have chosen a simple condition which is sufficient for our needs.

Proof.

Step 1. Suppose that $\boldsymbol{u}$ is not defined on $[0,1]$ and let $T_+ \leq 1$ be its final time of existence. In Step 2. we will prove (A.2) for all $t < T_+$. In particular, by the lower weak semi-continuity of the norm, this shows that

$$\sup_{t\in[0,T_+)} \|\boldsymbol{u}_n - \boldsymbol{v}\|_{\dot{H}^1\times L^2} \leq \delta.$$

By local well-posedness in the energy space and compactness of $\{v(t) : t \in [0,1]\}$, if $\delta > 0$ is small enough, there exists $\tau > 0$ such that the solution corresponding to the initial data $\boldsymbol{u}_n(t)$ is defined at least on the interval $(-\tau, \tau)$. This means that $\boldsymbol{u}$ cannot blow up at T_+, and so it is defined for $t \in [0,1]$.

If δ is chosen small enough, depending on $v(1)$, then by the Cauchy theory the solutions $\boldsymbol{u}_n$ exist on an interval $(1-t', 1+t')$ for some $t' > 0$. By eventually choosing t' smaller, we can assume that $\boldsymbol{u}$ also exists on $(1-t', 1+t')$. Hence, by repeating the same procedure we obtain weak convergence also for $t = 1$.

Step 2. Let $t < T_+$. In order to prove (A.2), it is sufficient to show that any subsequence of $\boldsymbol{u}_n$ (which we will still denote $\boldsymbol{u}_n$) admits a subsequence such that the required convergence takes place. By the result of Bahouri-Gérard a subsequence of $\boldsymbol{u}_n(0)$ admits a profile decomposition such that the first profile is $\boldsymbol{U}_{\text{LIN}}^1(t) = S(t)(u_0, u_1)$ (corresponding to parameters $t_{1,n} = 0$, $\lambda_{1,n} = 1$). By the triangle inequality $\|\boldsymbol{u}_n - (u_0, u_1)\|_{\dot{H}^1\times L^2} \leq 2\delta$, so all the other profiles are small, in particular they are global and disperse. By definition of T_+ the assumptions of Proposition 2.8 in [23] (which is a version of [3, Main Theorem] for the focusing equation) are satisfied for $\theta_n = t$, in particular formula (2.22) from [23] yields:

$$\boldsymbol{u}_n(t) = \boldsymbol{u}(t) + \sum_{j=2}^{J} \boldsymbol{U}_n^j(t) + \boldsymbol{w}_n^J(t) + \boldsymbol{r}_n^J(t).$$

Here, $\boldsymbol{w}_n^J(t) = S(t)\boldsymbol{w}_n^J(0) \rightharpoonup 0$ as $n \to +\infty$ (indeed, $\boldsymbol{w}_n^J(0) \rightharpoonup 0$ for $J > 1$ by definition of the profiles, and $S(t)$ is a bounded linear operator). By Lemma A.3 below also $U_n^j(t) \rightharpoonup 0$ when $j > 1$, which finishes the proof.

□

Lemma A.3. *Let* $\boldsymbol{U}$ *be a solution of equation* (NLW) *such that* $\|\boldsymbol{U}\|_{\dot{H}^1\times L^2}$ *is small. Let* t_n, λ_n *be a sequence of parameters such that one of the following holds:*

1. $t_n = 0$ *and* $\lambda_n \to 0$,
2. $t_n = 0$ *and* $\lambda_n \to +\infty$,
3. $t_n/\lambda_n \to +\infty$,
4. $t_n/\lambda_n \to -\infty$.

Fix $t \in \mathbb{R}$ *and define*

$$\boldsymbol{U}_n(x) = \Big(\frac{1}{\lambda_n^{3/2}}U^j\big(\frac{t-t_n}{\lambda_n}, \frac{x}{\lambda_n}\big), \frac{1}{\lambda_n^{5/2}}\partial_t U^j\big(\frac{t-t_n}{\lambda_n}\big)\Big).$$

Then $\boldsymbol{U}_n \rightharpoonup 0$ *in* $\dot{H}^1 \times L^2$.

Proof. Again it is sufficient to show this for a subsequence of any subsequence. Thus we can assume that $\frac{t-t_n}{\lambda_n} \to t_0 \in [-\infty, +\infty]$.

Suppose first that t_0 is a finite number. Extracting again a subsequence we can assume that $\lambda_n \to \lambda_0 \in [0, +\infty]$. If λ_0 was a strictly positive finite number, we would obtain that also t_n has a finite limit, which is impossible. Thus $\lambda_n \to 0$ or $\lambda_n \to +\infty$, and in both cases we get our conclusion.

In the case $\frac{t-t_n}{\lambda_n} \to \pm\infty$ we have dispersion, so $\|\boldsymbol{U}_n - (S(\tau_n)\boldsymbol{V})_{\lambda_n}\|_{\dot{H}^1\times L^2} \to 0$, and it is well known that $(S(\tau_n)\boldsymbol{V})_{\lambda_n} \rightharpoonup 0$ when $\tau_n \to \pm\infty$ and λ_n is any sequence (in the case of space dimension $N = 5$ this follows for example from the strong Huyghens principle). □

B Local theory in higher regularity

In this section we use the energy method to prove two results about preservation of regularity.

B.1 Energy estimates in $X^1 \times H^1$

Recall that we denote $X^s := \dot{H}^{s+1} \cap \dot{H}^1$. We have classical *energy estimates* for the linear wave equation:

Lemma B.1. *Let* $s \in \mathbb{N}$. *Let* $I = [0, T_0]$ *be a time interval,* $g \in C(I, H^s)$ *and* $(u_0, u_1) \in X^s \times H^s$. *Then the Cauchy problem*

$$\begin{cases} \partial_{tt}u - \Delta u = g, \\ (u(0), \partial_t u(0)) = (u_0, u_1) \end{cases}$$

has a unique solution $(u, \partial_t u) \in C(I, X^s \times H^s)$ *and for all* $t \in I$ *there holds*

$$\|(u, \partial_t u)\|_{X^s\times H^s} \leq \|(u_0, u_1)\|_{X^s\times H^s} + \int_0^t \|g(\tau)\|_{H^s}\,\mathrm{d}\tau. \tag{B.1}$$

For a proof of a more general result one can consult for example [2, Theorem 4.4]. Using finite speed of propagation and Sobolev Extension Theorem on each time slice we get a localised version of the energy estimate:

$$\|(u, \partial_t u)\|_{X^s\times H^s(B(0,\rho))} \lesssim \|(u_0, u_1)\|_{X^s\times H^s(B(0,\rho+t))} + \int_0^t \|g(\tau)\|_{H^s(B(0,\rho+\tau)}\,\mathrm{d}\tau \tag{B.2}$$

Now we use the case $s = 1$ to prove energy estimates in $X^1 \times H^1$ for (NLW).

Proposition B.2. *For all $M_0 > 0$ there exists $T_0 = T_0(M_0) > 0$ such that the following is true. Let $(u_0, u_1) \in X^1 \times H^1$ with $\|(u_0, u_1)\|_{X^1\times H^1} \leq M_0$. Then the Cauchy problem:*

$$\begin{cases} \partial_{tt}u - \Delta u = f(u), \\ (u(0), \partial_t u(0)) = (u_0, u_1) \end{cases}$$

has a unique solution $(u, \partial_t u) \in C([0,T_0], X^1 \times H^1)$ and this solution verifies

$$\sup_{t\in[0,T_0]} \|(u(t), \partial_t u(t))\|_{X^1\times H^1} \leq 2\|(u_0, u_1)\|_{X^1\times H^1}.$$

Moreover, let u_{LIN} denote the solution of the free wave equation for the same initial data (u_0, u_1). Then

$$\sup_{t\in[0,T_0]} \|(u(t), \partial_t u(t)) - (u_{\text{LIN}}(t), \partial_t u_{\text{LIN}}L(t))\|_{X^1\times H^1} \lesssim f\big(\|(u_0, u_1)\|_{X^1\times H^1}\big). \tag{B.3}$$

This will follow easily from the following lemma.

Lemma B.3. *Let $u, v \in X^1$. Then*

$$\|f(u)\|_{H^1} \leq Cf(\|u\|_{X^1}), \tag{B.4}$$
$$\|f(u) - f(v)\|_{H^1} \leq C\|u - v\|_{X^1} \cdot \big(f'(\|u\|_{X^1}) + f'(\|v\|_{X^1})\big). \tag{B.5}$$

Proof. We have $\|f(u)\|_{L^2} = \|u\|^{7/3}_{L^{14/3}} \lesssim f(\|u\|_{X^1})$ from the Sobolev imbedding. By Hölder inequality,

$$\|\nabla f(u)\|_{\dot{H}^1} = \|\nabla u \cdot f'(u)\|_{L^2} \lesssim \|\nabla u\|_{L^{10/3}} \cdot \|f'(u)\|_{L^5} \lesssim \|u\|_{X^1} \cdot \|u\|^{4/3}_{L^{20/3}} \lesssim f(\|u\|_{X^1}),$$

again by Sobolev imbedding. This proves (B.4).

To prove (B.5), we write $|f(u) - f(v)| \lesssim |u - v|(f'(u) + f'(v))$, hence

$$\begin{aligned} \|f(u) - f(v)\|_{L^2} &\lesssim \|u - v\|_{L^{14/3}} \cdot \|f'(u) + f'(v)\|_{L^{7/2}} \lesssim \|u - v\|_{L^{14/3}} \cdot \big(\|u\|^{4/3}_{L^{14/3}} + \|v\|^{4/3}_{L^{14/3}}\big) \\ &\lesssim \|u - v\|_{X^1} \cdot \big(f'(\|u\|_{X^1}) + f'(\|v\|_{X^1})\big). \end{aligned}$$

Finally,

$$|\nabla f(u) - \nabla f(v)| \lesssim |\nabla u - \nabla v|(f'(u) + f'(v)) + |u - v|(|\nabla u| + |\nabla v|)(|f''(u)| + |f''(v)|),$$

and it suffices to notice that

$$\begin{aligned} \||\nabla u - \nabla v|(f'(u) + f'(v))\|_{L^2} &\lesssim \|\nabla u - \nabla v\|_{L^{10/3}} \cdot \|f'(u) + f'(v)\|_{L^5} \\ &\lesssim \|u - v\|_{X^1} \cdot \big(f'(\|u\|_{X^1}) + f'(\|v\|_{X^1})\big) \end{aligned}$$

and

$$\begin{aligned} &\big\||u - v|(|\nabla u| + |\nabla v|)(|f''(u)| + |f''(v)|)\big\|_{L^2} \\ &\lesssim \|u - v\|_{L^{10}} \cdot (\|\nabla u\|_{L^{10/3}} + \|\nabla v\|_{L^{10/3}}) \cdot \big(\|f''(u)\|_{L^{10}} + \|f''(v)\|_{L^{10}}\big) \\ &\lesssim \|u - v\|_{X^1} \cdot \big(f'(\|u\|_{X^1}) + f'(\|v\|_{X^1})\big). \end{aligned}$$

□

Proof of Proposition B.2. Let B denote the ball of centre 0 and radius $2\|(u_0,u_1)\|_{X^1\times H^1}$ in the space $X^1\times H^1$. Given $(u,\partial_t u)\in C([0,T],B)$, let $\widetilde{u}=\Phi(u)$ denote the solution of the Cauchy problem

$$\begin{cases}\partial_{tt}\widetilde{u}-\Delta\widetilde{u}=f(u),\\(\widetilde{u}(0),\partial_t\widetilde{u}(0))=(u_0,u_1)\end{cases}$$

It follows from Lemma (B.4) and (B.1) that if $T\leq\frac{M_0}{Cf(2M_0)}$, then $(\widetilde{u},\partial_t\widetilde{u})\in C([0,T],B)$. It follows from (B.5) and (B.1) that if $T\leq\frac{1}{4Cf'(2M_0)}$, then Φ is a contraction, so it has a unique fixed point, which is the desired solution.

The function $v:=u-u_{\mathrm{LIN}}$ solves the Cauchy problem

$$\begin{cases}\partial_{tt}v-\Delta v=f(u),\\(v(0),\partial_t v(0))=0,\end{cases}$$

so (B.3) follows from (B.1). □

B.2 Persistence of $X^1\times H^1$ regularity

We recall the classical Strichartz inequality:

Lemma B.4. *[33] Let $I=[0,T_0]$ be a time interval, $g\in C(I,L^2)$ and $(u_0,u_1)\in\dot{H}^1\times L^2$. Let $\boldsymbol{u}$ be the solution of the Cauchy problem*

$$\begin{cases}\partial_{tt}u-\Delta u=g,\\(u(0),\partial_t u(0))=(u_0,u_1).\end{cases}$$

Then

$$\|u\|_{L^{7/3}(I;L^{14/3})}\lesssim\|(u_0,u_1)\|_{\dot{H}^1\times L^2}+\|g\|_{L^1(I;L^2)},$$

with a constant independent of I.

From the local theory of (NLW) in the critical space we know that if $\boldsymbol{u}\in C((T_-,T_+);\dot{H}^1\times L^2)$ is a solution of (NLW) and $I=[T_1,T_2]\subset(T_-,T_+)$, then

$$\|u\|_{L^{7/3}(I;L^{14/3})}<+\infty. \tag{B.6}$$

Proposition B.5. *Suppose that $0\in I=[T_1,T_2]\subset(T_-,T_+)$ and that $(u_0,u_1)\in X^1\times H^1$. Then $\boldsymbol{u}\in C(I,X^1\times H^1)$.*

Proof. The proof is classical, see for example [9, Chapter 5] for more general results in the case of NLS.

We consider positive times. The proof for negative times is the same. Let T_* be the maximal time of existence of $\boldsymbol{u}$ in $X^1\times H^1$. Suppose that $T_*<T_+$. From Proposition B.2 it follows that

$$\lim_{t\to T_*}\|\boldsymbol{u}\|_{X^1\times H^1}=+\infty. \tag{B.7}$$

Consider the time interval $I=[T_*-\tau,T_*]$. Derivating (NLW) once and using Lemma B.4 we get

$$\|\nabla u\|_{L^{7/3}(I;L^{14/3})}\leq C\|(u(T_*-\tau),\partial_t u(T_*-\tau))\|_{X^1\times H^1}+C\|\nabla(f(u))\|_{L^1(I;L^2)}, \tag{B.8}$$

with C independent of τ. From Hölder inequality we have

$$\|\nabla(f(u))\|_{L^1(I;L^2)}\leq\|\nabla u\|_{L^{7/3}(I;L^{14/3})}\cdot f'\big(\|u\|_{L^{7/3}(I;L^{14/3})}\big).$$

By (B.6), the last term is arbitrarily small when $\tau \to 0^+$, so for τ small enough the second term on the right hand side of (B.8) can be absorbed by the left hand side, which implies $\|\nabla u\|_{L^{7/3}(I;L^{14/3})} < +\infty$ and $\|\nabla(f(u))\|_{L^1(I;L^2)} < +\infty$. This is in contradiction with (B.7), because of the energy estimate (B.1). □

B.3 Propagation of regularity around a non-degenerate point

Proposition B.6. *Let* $(u_0, u_1) \in X^4 \times H^4$ *such that* $u_0(0) > 0$. *Let* $(u, \partial_t u) \in C([0, T_0]; X^1 \times H^1)$ *be the solution of the Cauchy problem:*

$$\begin{cases} \partial_{tt} u - \Delta u = f(u), \\ (u(0), \partial_t u(0)) = (u_0, u_1), \end{cases}$$

constructed in Proposition B.2. There exists $\tau, \rho > 0$ *such that* $(u, \partial_t u)$ *satisfies:*

$$\big(\chi(\frac{\cdot}{\rho})u, \chi(\frac{\cdot}{\rho})\partial_t u\big) \in C([0, \tau]; X^4 \times H^4)$$

(where χ *is a standard regular cut-off function).*

Proof. Denote $v_0 := u_0(0) > 0$ and introduce an auxiliary function $\widetilde{f} \in C^\infty$, $\widetilde{f}(u) = f(u)$ when $u \geq v_0/2$, $f(u) = 0$ when $u \leq 0$. Using Faà di Bruno formula one can prove an analog of Lemma B.1:

$$\|\widetilde{f}(u)\|_{H^4} \leq C(\|u\|_{X^4}),$$
$$\|\widetilde{f}(u) - \widetilde{f}(v)\|_{H^4} \leq \|u - v\|_{X^4} \cdot C(\|u\|_{X^4} + \|v\|_{X^4}),$$

where $C : \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous function. The same procedure as in the proof of Proposition B.2 leads to the conlusion that there exists $\tau > 0$ such that the Cauchy problem:

$$\begin{cases} \partial_{tt} \widetilde{u} - \Delta \widetilde{u} = \widetilde{f}(\widetilde{u}), \\ (\widetilde{u}(0), \partial_t \widetilde{u}(0)) = (u_0, u_1) \end{cases}$$

has a solution $(\widetilde{u}, \partial_t \widetilde{u}) \in C([0, \tau], X^4 \times H^4)$. By continuity and Schauder estimates, if we take τ and ρ sufficiently small, we have $\widetilde{u}(t, x) > \frac{1}{2}v_0$ for $|x| \leq 4\rho$ and $0 \leq t \leq \tau$. We may assume that $\tau \leq 2\rho$. Consider $v = u - \widetilde{u}$. We will prove that $v = 0$ when $0 \leq t \leq \tau$ and $|x| \leq 2\rho$, which will finish the proof. The function v solves the Cauchy problem:

$$\begin{cases} \partial_{tt} v - \Delta v = f(u) - \widetilde{f}(\widetilde{u}), \\ (v(0), \partial_t v(0)) = 0. \end{cases}$$

We run the localized energy estimate (B.2) for $|x| \leq 2\rho + |t - \tau|$. We suppose that $\tau \leq 2\rho$, so $|x| \leq 4\rho$, which means that $\|f(u) - \widetilde{f}(\widetilde{u})\|_{H^1} = \|f(u) - f(\widetilde{u})\|_{H^1} \lesssim \|u - \widetilde{u}\|_{X^1}$ (the norm is taken in the ball $B(0, 2\rho + |t - \tau|)$. From (B.2) and Gronwall inequality we deduce that $u = \widetilde{u}$ when $|x| \leq 2\rho + |t - \tau|$, in particular when $|x| \leq 2\rho$. □

B.4 Short-time asymptotics in the case $(u_0, u_1) = (p|x|^\beta, 0)$.

Let $(u, \partial_t u)$ denote the solution of (NLW) corresponding to the intial data

$$(u_0, u_1) = \big(\chi(\frac{\cdot}{\rho})p|x|^\beta, 0\big),$$

where $p, \rho > 0$ and $\beta > \frac{5}{2}$ are constants and χ is a standard cut-off function. Let $(u_{\text{LIN}}, \partial_t u_{\text{LIN}})$ denote the solution of the free wave equation corresponding to the same initial data.

Proposition B.7. *Let* $q = \frac{(\beta+1)(\beta+3)}{3}p$*. There exist* $T_0 > 0$ *and a constant* $C > 0$ *such that for* $0 \leq t \leq T_0$ *and* $|x| \leq \frac{1}{2}t$ *there holds*

$$|u_{\mathrm{LIN}}(t,x) - qt^\beta| \leq Ct^{\beta-2}|x|^2.$$

Proof. Define

$$w(y) := \fint_{\partial B(0,1)} p|\omega + ye_1|^\beta \, \mathrm{d}\sigma(\omega), \qquad -\frac{1}{\sqrt{2}} < y < \frac{1}{\sqrt{2}},$$

where $B(0,1)$ denote the unit ball in $\mathbb{R}^5$, $\mathrm{d}\sigma$ is the surface measure on the unit sphere and $e_1 = (1,0,0,0,0)$. Notice that

$$|\omega + ye_1|^\beta = (1 - \omega_1^2 + (y+\omega_1)^2)^{\beta/2} = (1+\omega_1^2)^{\beta/2} \cdot \big(1 + y\frac{2\omega_1}{1+\omega_1^2}\big)^{\beta/2}$$

can be developped in a power series in y which converges uniformly for $-\frac{1}{\sqrt{2}} < y < \frac{1}{\sqrt{2}}$. Hence, w is an analytic function. It is also symetric, so it is in fact analytic in y^2,

$$w(y) = \widetilde{w}(y^2), \qquad \widetilde{w}(z) \text{ analytic for } |z| < \frac{1}{2}.$$

We have $\widetilde{w}(0) = w(0) = p$.

The representation formula for solutions of the free wave equation, see for example [29, p. 77], yields

$$u_{\mathrm{LIN}}(t,x) = \frac{1}{3}\big(\frac{\partial}{\partial t}\big)\big(\frac{1}{t}\frac{\partial}{\partial t}\big)\big(t^3 \fint_{\partial B(x,t)} p|y|^\beta \, \mathrm{d}\sigma(y)\big).$$

A change of variables shows that for $|x| < \frac{1}{\sqrt{2}}t$ and t sufficiently small we have

$$u_{\mathrm{LIN}}(t,x) = \frac{1}{3}\big(\frac{\partial}{\partial t}\big)\big(\frac{1}{t}\frac{\partial}{\partial t}\big)\big(t^3 \cdot t^\beta \widetilde{w}\big(\frac{|x|^2}{t^2}\big)\big) = t^\beta \widetilde{w}_1\big(\frac{|x|^2}{t^2}\big),$$

where $\widetilde{w}_1(z)$ is analytic for $|z| < \frac{1}{2}$. It is easily seen that $\widetilde{w}_1(0) = \frac{(\beta+1)(\beta+3)}{3}p = q$ (all the terms coming from differentiating $\widetilde{w}$ vanish at $z = 0$). Hence, there exists a constant C such that $|\widetilde{w}_1(z) - q| \leq C|z|$ for $|z| \leq \frac{1}{4}$, and the conclusion follows. □

Proposition B.8. *For* t *small enough there holds*

$$\|u - u_{\mathrm{LIN}}\|_{X^1(|x| \leq \frac{1}{2}t)} \lesssim t^{\frac{7}{3}\beta + \frac{7}{6}}.$$

Proof. From (B.3) and finite speed of propagation we obtain

$$\|u - u_{\mathrm{LIN}}\|_{X^1(|x| \leq \frac{1}{2}t)} \lesssim f(\|(u_0,u_1)\|)_{X^1 \times H^1(|x| \leq \frac{3}{2}t)}.$$

We have

$$\|(u_0,u_1)\|^2_{X^1 \times H^1(|x| \leq \frac{3}{2}t)} \sim \int_0^{\frac{3}{2}t} (r^{\beta-2})^2 r^4 \, \mathrm{d}r \sim t^{2\beta+1},$$

and the conclusion follows. □

Chapter 2

Construction of two-bubble solutions for energy-critical wave equations

Abstract

We construct pure two-bubbles for some energy-critical wave equations, that is solutions which in one time direction approach a superposition of two stationary states both centered at the origin, but asymptotically decoupled in scale. Our solution exists globally, with one bubble at a fixed scale and the other concentrating in infinite time, with an error tending to 0 in the energy space. We treat the cases of the power nonlinearity in space dimension 6, the radial Yang-Mills equation and the equivariant wave map equation with equivariance class $k \geq 3$. The concentrating speed of the second bubble is exponential for the first two models and a power function in the last case.

1 Introduction

1.1 Energy critical NLW

We consider the energy critical wave equation in space dimension $N = 6$:

$$\begin{cases} (\partial_t^2 - \Delta)u(t,x) = |u(t,x)| \cdot u(t,x), & (t,x) \in \mathbb{R}\times\mathbb{R}^6, \\ (u(t_0,x), \partial_t u(t_0,x)) = (u_0(x), \dot u_0(x)). \end{cases} \tag{1.1}$$

The *energy functional* associated with this equation is defined for $\boldsymbol{u}_0 = (u_0, \dot u_0) \in \mathcal{E} := \dot H^1(\mathbb{R}^6)\times L^2(\mathbb{R}^6)$ by the formula

$$E(\boldsymbol{u}_0) := \int \frac12 |\dot u_0|^2 + \frac12|\nabla u_0|^2 - F(u_0)\,\mathrm{d}x,$$

where $F(u_0) := \frac13|u_0|^3$. Note that $E(\boldsymbol{u}_0)$ is well-defined due to the Sobolev Embedding Theorem. The differential of E is $\mathrm{D}E(\boldsymbol{u}_0) = (-\Delta u_0 - f(u_0), \dot u_0)$, where $f(u_0) = |u_0|\cdot u_0$, hence we can rewrite equation (1.1) as

$$\begin{cases} \partial_t \boldsymbol{u}(t) = J\circ \mathrm{D}E(\boldsymbol{u}(t)), \\ \boldsymbol{u}(t_0) = \boldsymbol{u}_0 \in \mathcal{E}. \end{cases} \tag{1.2}$$

Here, $J := \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0\end{pmatrix}$ is the natural symplectic structure.

Equation (1.1) is locally well-posed in the space $\mathcal{E}$, see for example Ginibre, Soffer and Velo [32], Shatah and Struwe [84] (the defocusing case), as well as a complete review of the Cauchy theory in Kenig and Merle [47] (for $N \in \{3,4,5\}$) and Bulut, Czubak, Li, Pavlović and Zhang [8] (for $N \geq 6$). By "well-posed" we mean that for any initial data $\boldsymbol{u}_0 \in \mathcal{E}$ there exists $\tau > 0$ and a unique solution in some subspace of $C([t_0-\tau, t_0+\tau];\mathcal{E})$, and that this solution is continuous with respect to the inital data. By standard arguments, there exists a maximal time of existence (T_-, T_+), $-\infty \leq T_- < t_0 < T_+ \leq +\infty$, and a unique solution $\boldsymbol{u} \in C((T_-,T_+);\mathcal{E})$. If $T_+ < +\infty$, then $\boldsymbol{u}(t)$ leaves every compact subset of $\mathcal{E}$ as t approaches T_+. A crucial property of the solutions of (1.1) is that the energy E is a conservation law.

In this paper we always assume that the initial data are radially symmetric. This symmetry is preserved by the flow.

For functions $v \in \dot H^1$, $\dot v \in L^2$, $\boldsymbol{v} = (v, \dot v) \in \mathcal{E}$ and $\lambda > 0$, we denote

$$v_\lambda(x) := \frac{1}{\lambda^2} v\Big(\frac{x}{\lambda}\Big), \qquad \dot v_{\underline{\lambda}}(x) := \frac{1}{\lambda^3}\dot v\Big(\frac{x}{\lambda}\Big), \qquad \boldsymbol{v}_\lambda(x) := \big(v_\lambda, \dot v_{\underline{\lambda}}\big).$$

A change of variables shows that

$$E\big((\boldsymbol{u}_0)_\lambda\big) = E(\boldsymbol{u}_0).$$

Equation (1.1) is invariant under the same scaling: if $\boldsymbol{u}(t) = (u(t), \dot u(t))$ is a solution of (1.1) and $\lambda > 0$, then $t \mapsto \boldsymbol{u}\big((t-t_0)/\lambda\big)_\lambda$ is also a solution with initial data $(\boldsymbol{u}_0)_\lambda$ at time $t = 0$. This is why equation (1.1) is called *energy-critical*.

A fundamental object in the study of (1.1) is the family of stationary solutions $\boldsymbol{u}(t) \equiv \pm\boldsymbol{W}_\lambda = (\pm W_\lambda, 0)$, where

$$W(x) = \Big(1 + \frac{|x|^2}{24}\Big)^{-2}.$$

The functions W_λ are called *ground states* or *bubbles* (of energy). They are the only radially symmetric solutions and, up to translation, the only positive solutions of the critical elliptic problem

$$-\Delta u - f(u) = 0.$$

The ground states achieve the optimal constant in the critical Sobolev inequality, which was proved by Aubin [1] and Talenti [90]. They are the "mountain passes" for the potential energy.

Kenig and Merle [47] exhibited the special role of the ground states $\boldsymbol{W}_\lambda$ as the *threshold elements* for nonlinear dynamics of the solutions of (1.1) in space dimensions $N = 3, 4, 5$, which is believed to be a general feature of dispersive equations (the so-called *Threshold Conjecture*). Another major problem in the field is the *Soliton Resolution Conjecture*, which predicts that a bounded (in an appropriate sense) solution decomposes asymptotically into a sum of energy bubbles at different scales and a radiation term (a solution of the linear wave equation). This was proved for the radial energy-critical wave equation in dimension $N = 3$ by Duyckaerts, Kenig and Merle [26], following the earlier work of the same authors [24], where such a decomposition was proved only for a sequence of times (this last result was generalized to any odd dimension by Rodriguez [81]).

It is natural to examine the dynamics of the solutions of (1.1) in a neighborhood (in the energy space) of the family of the ground states. In dimension $N = 3$ the was done by Krieger, Schlag and Tataru [53], who showed that such solutions can blow up in finite time (by concentration of the bubble), see also [20], [50], [19], [36], [40] for related results.

In view of the rich dynamics in a neighborhood of one bubble, it was expected that solutions behaving asymptotically as a superposition of many (at least two) bubbles exist, in other words that the result of [26] is essentially optimal. We prove that it is the case when $N = 6$:

Theorem 1. *There exists a solution* $\boldsymbol{u} : (-\infty, T_0] \to \mathcal{E}$ *of* (1.1) *such that*

$$\lim_{t\to-\infty} \|\boldsymbol{u}(t) - (\boldsymbol{W} + \boldsymbol{W}_{\frac{1}{\kappa}\mathrm{e}^{-\kappa|t|}})\|_{\mathcal{E}} = 0, \qquad \textit{with } \kappa := \sqrt{\frac{5}{4}}.$$

Remark 1.1. More precisely, we will prove that

$$\left\|\boldsymbol{u}(t) - \left(W + W_{\frac{1}{\kappa}\mathrm{e}^{-\kappa|t|}}, -\mathrm{e}^{-\kappa|t|}\Lambda W_{\underline{\frac{1}{\kappa}\mathrm{e}^{-\kappa|t|}}}\right)\right\|_{\mathcal{E}} \leq C_1 \cdot \mathrm{e}^{-\frac{3}{2}\kappa|t|},$$

where $\Lambda W := -\frac{\partial}{\partial\lambda} W_\lambda\big|_{\lambda=1}$ and $C_1 > 0$ is a constant.

Remark 1.2. We construct here *pure* two-bubbles, that is the solution approaches a superposition of two stationary states, with no energy transformed into radiation. By the conservation of energy and the decoupling of the two bubbles, we necessarily have $E(\boldsymbol{u}(t)) = 2E(\boldsymbol{W})$. Pure one-bubbles cannot concentrate and are completely classified, see [27].

Remark 1.3. It was proved in [41], in any dimension $N \geq 3$, that there exist no solutions $\boldsymbol{u}(t) : [t_0, T_+) \to \mathcal{E}$ of (1.1) such that $\|\boldsymbol{u}(t) - (\boldsymbol{W}_{\mu(t)} - \boldsymbol{W}_{\lambda(t)})\|_{\mathcal{E}} \to 0$ with $\lambda(t) \ll \mu(t)$ as $t \to T_+ \leq +\infty$.

Remark 1.4. In any dimension $N > 6$ one can expect an analogous result with concentration rate $\lambda(t) \sim |t|^{-\frac{4}{N-6}}$.

Remark 1.5. In the context of the harmonic map heat flow, Topping [93] proved the existence of towers of bubbles for a well chosen target manifold, see also a non-existence result of van der Hout [37].

Let us resume the overall strategy of the proof, which is based on the previous paper of the author [40].

In Section 2 we construct an appropriate approximate solution $\varphi(t)$. We present first a formal computation which allows to predict the concentration rate and explains why the proof fails in dimension $N \in \{3, 4, 5\}$. It highlights also the role of the *strong interaction* between the two bubbles (by "strong" we mean "significantly altering the dynamics"; [62] provides an example of this phenomenon in a different context). Then we give a precise definition of the approximate solution and prove bounds on its error.

In Section 3 we build a sequence $\boldsymbol{u}_n : [t_n, T_0] \to \mathcal{E}$ of solutions of (1.1) with $t_n \to -\infty$ and $\boldsymbol{u}_n(t)$ close to a two-bubble solution for $t \in [t_n, T_0]$. Taking a weak limit finishes the proof. This type of argument goes back to the works of Merle [64] and Martel [56]. The heart of the analysis is to obtain uniform energy bounds for the sequence $\boldsymbol{u}_n$. To this end we follow the approach of Raphaël and Szeftel [78], that is we prove bootstrap estimates involving an energy functional with a virial-type correction term. This correction is designed to cancel some terms related to the concentration of the bubble $\boldsymbol{W}_{\lambda(t)}$. It has to be localized in an appropriate way, so that it does not "see" the other bubble. Finally, in order to deal with the linear instabilities of the flow, we use a classical topological ("shooting") argument.

1.2 Critical wave maps

We consider the wave map equation from the 2+1-dimensional Minkowski space (the energy-critical case) to $\mathbb{S}^2$. We will consider solutions with k-equivariant symmetry, in which case the problem is reduced to the following scalar equation:

$$\begin{cases} \partial_t^2 u(t,r) = \partial_r^2 u(t,r) + \frac{1}{r}\partial_r u(t,r) - \frac{k^2}{2r^2}\sin(2u(t,r)), \\ (u(t_0,r), \partial_t u(t_0,r)) = (u_0(r), \dot u_0(r)), \qquad t, t_0 \in \mathbb{R},\ r \in (0,+\infty). \end{cases} \tag{1.3}$$

For a presentation of the geometric content of this equation, one can consult [85]. Here we will regard (1.3) as a scalar semilinear problem.

We define the space $\mathcal{H}$ as the completion of $C_0^\infty((0,+\infty))$ for the norm

$$\|v\|_{\mathcal{H}}^2 := 2\pi\int_0^{+\infty}\big(|\partial_r v(r)|^2 + |\frac{k}{r}v(r)|^2\big)\,r\mathrm{d}r.$$

We will work in the energy space $\mathcal{E} := \mathcal{H} \times L^2$. Equation (1.3) can be written in the form (1.2) with the energy functional E defined for $\boldsymbol{u}_0 = (u_0, \dot u_0) \in \mathcal{E}$ by the formula

$$E(\boldsymbol{u}_0) := \pi\int_0^{+\infty}\big((\dot u_0)^2 + (\partial_r u_0)^2 + \frac{k^2}{r^2}(\sin(u))^2\big)\,r\mathrm{d}r.$$

The Cauchy theory in the energy space has been established by Shatah and Tahvildar-Zadeh [86]. Note that $u_0 \in \mathcal{H}$ forces $\lim_{r\to+\infty} u_0(r) = 0$, but we could just as well consider states of finite energy such that $\lim_{r\to+\infty} u_0(r) = \pi$, see [17, 16] for details.

The stationary solutions $W_\lambda(r) := 2\arctan\big(\big(\frac{r}{\lambda}\big)^k\big)$ play a fundamental role in the study of (1.3). They are the harmonic maps of topological degree k. We will write $W(r) := W_1(r) = 2\arctan(r^k)$ and $\Lambda W(r) := -\frac{\partial}{\partial\lambda}W_\lambda\big|_{\lambda=1} = \frac{2k}{r^k+r^{-k}}$. Note that $W \notin \mathcal{H}$ precisely because of the fact that $W(r) \to \pi$ as $r \to +\infty$.

The possibility of concentration of a harmonic map at the origin was first observed numerically by Bizoń, Chmaj and Tabor [4]. Struwe [89] proved that if the blow-up occurs, then $\boldsymbol{W}$ is the blow-up profile (for a sequence of times). The dynamics in a neighborhood of a harmonic map was studied by Krieger, Schlag and Tataru [52], who constructed blow-up solutions in the energy space with the concentration rate $\lambda(t) \sim (T_+ - t)^{1+\nu}$ for all $\nu > \frac{1}{2}$. This behavior is expected to be highly unstable. Rodnianski and Sterbenz [80] constructed stable blow-up solutions, giving the first (partial) rigorous explanation of the surprising numerical results mentioned above. In the case $k = 1$, Côte [14] proved that any solution decomposes, for a sequence of times tending to the final (finite or inifinite) time of existence, as a sum of a finite number of harmonic maps at different scales and a radiation term. A generalization of this result, including all the cases considered in this paper, was recently obtained by Jia and Kenig [42]. Motivated by these works, we prove the following result.

Theorem 2. *Fix $k > 2$. There exists a solution $\boldsymbol{u} : (-\infty, T_0] \to \mathcal{E}$ of (1.3) such that*

$$\lim_{t\to-\infty} \|\boldsymbol{u}(t) - (-\boldsymbol{W} + \boldsymbol{W}_{\frac{k-2}{\kappa}(\kappa|t|)^{-\frac{2}{k-2}}})\|_{\mathcal{E}} = 0, \qquad \text{with } \kappa := \frac{k-2}{2}\Big(\frac{8k}{\pi}\sin\big(\frac{\pi}{k}\big)\Big)^{\frac{1}{k}}.$$

Remark 1.6. More precisely, we will prove that

$$\big\|\boldsymbol{u}(t) - \big(-W + W_{\frac{k-2}{\kappa}(\kappa|t|)^{-\frac{2}{k-2}}}, -(\kappa|t|)^{-\frac{k}{k-2}}\Lambda W_{\underline{\frac{k-2}{\kappa}(\kappa|t|)^{-\frac{2}{k-2}}}}\big)\big\|_{\mathcal{E}} \le C_1 \cdot |t|^{-\frac{k+1}{k-2}},$$

where $\Lambda W := -\frac{\partial}{\partial\lambda} W_\lambda\big|_{\lambda=1}$ and $C_1 > 0$ is a constant.

Remark 1.7. The constructed solution is a *pure* two-bubble, hence by the conservation of energy $E(\boldsymbol{u}(t)) = 2E(W)$, and it is clear that it has the homotopy degree 0. In the case of equivariant class $k = 1$, Côte, Kenig, Lawrie and Schlag [16] showed that any degree 0 initial data of energy $< 2E(W)$ leads to dispersion (the proof is expected to generalize to all equivariance classes). Theorem 2 gives the first example of a non-dispersive solution at the threshold energy.

Note that pure two-bubbles of homotopy degree $2k$ (hence of type bubble-bubble and not bubble-antibubble) do not exist because the energy of such a map has to be $> 2E(\boldsymbol{W})$. This is similar to the case of opposite signs for (1.1), see Remark 1.3.

Remark 1.8. I believe that the proof can by adapted to deal with a more general equation $\partial_t^2 u = \partial_r^2 u + \frac{1}{r}\partial_r u - \frac{1}{r^2}(gg')(u)$ with g satisfying the assumptions of [17] and $g'(0) \in \{3, 4, 5, \ldots\}$.

1.3 Critical Yang-Mills

Finally, we consider the radial Yang-Mills equation in dimension 4 (which is the energy-critical case):

$$\begin{cases} \partial_t^2 u(t,r) = \partial_r^2 u(t,r) + \dfrac{1}{r}\partial_r u(t,r) - \dfrac{4}{r^2}u(t,r)(1-u(t,r))\big(1-\dfrac{1}{2}u(t,r)\big), \\ (u(t_0,r), \partial_t u(t_0,r)) = (u_0(r), \dot u_0(r)), \qquad t, t_0 \in \mathbb{R},\ r \in (0,+\infty). \end{cases} \tag{1.4}$$

For a derivation of this equation and further comments, see for instance [10]. Equation (1.4) can be written in the form (1.2) with the energy functional E defined for $\boldsymbol{u}_0 = (u_0, \dot u_0) \in \mathcal{E}$ by the formula

$$E(\boldsymbol{u}_0) := \pi \int_0^{+\infty} \big((\dot u_0)^2 + (\partial_r u_0)^2 + \frac{1}{r^2}(u_0(2-u_0))^2\big)\, r\mathrm{d}r.$$

The stationary solutions of (1.4) are $W_\lambda(r) := \frac{2r^2}{\lambda^2+r^2}$. We denote $W(r) := W_1(r) = \frac{2r^2}{1+r^2}$ and $\Lambda W(r) := -\frac{\partial}{\partial\lambda}W_\lambda\big|_{\lambda=1} = \frac{4}{(r+r^{-1})^2}$.

Theorem 3. *There exists a solution* $\boldsymbol{u} : (-\infty, T_0] \to \mathcal{E}$ *of* (1.4) *such that*

$$\lim_{t\to-\infty} \|\boldsymbol{u}(t) - (-\boldsymbol{W} + \boldsymbol{W}_{\frac{1}{\kappa}\mathrm{e}^{-\kappa|t|}})\|_{\mathcal{E}} = 0, \qquad \textit{with } \kappa := 2\sqrt{3}.$$

Remark 1.9. More precisely, we will prove that

$$\big\|\boldsymbol{u}(t) - \big(-W + W_{\frac{1}{\kappa}\mathrm{e}^{-\kappa|t|}}, -\mathrm{e}^{-\kappa|t|}\Lambda W_{\underline{\frac{1}{\kappa}\mathrm{e}^{-\kappa|t|}}}\big)\big\|_{\mathcal{E}} \leq C_1 \cdot \mathrm{e}^{-\frac{3}{2}\kappa|t|},$$

where $\Lambda W := -\frac{\partial}{\partial\lambda}W_\lambda\big|_{\lambda=1}$ and $C_1 > 0$ is a constant.

Remark 1.10. The case of wave maps in the equivariance class $k = 2$ should be very similar.

Remark 1.11. The energy $2E(\boldsymbol{W})$ is the threshold energy for a non-dispersive behavior for solutions with topological degree 0, see [54].

1.4 Structure of the paper

In Sections 2 and 3 we give a detailed proof of Theorem 1. In Section 4 we treat the case of the Yang-Mills equation. We skip these parts of the proof where the arguments of Sections 2 and 3 are directly applicable. Section 5 is devoted to the wave maps equation. The main difference with respect to Section 4 is that the characteristic length of the concentrating bubble is now a power of $|t|$ and not an exponential. Nevertheless, large parts of the previous proofs extend to this case and are skipped. It is conceivable that one could propose a unified, more general framework of the proof, encompassing all the cases under consideration. Appendix A is devoted to some elements of the local Cauchy theory needed in the proofs.

1.5 Notation

The bracket $\langle\cdot,\cdot\rangle$ denotes the distributional pairing and the scalar product in the spaces L^2 and $L^2 \times L^2$.

For positive quantities m_1 and m_2 we write $m_1 \lesssim m_2$ if $m_1 \leq Cm_2$ for some constant $C > 0$ and $m_1 \sim m_2$ if $m_1 \lesssim m_2 \lesssim m_1$.

We denote χ a standard C^∞ cut-off function, that is $\chi(x) = 1$ for $|x| \leq 1$, $\chi(x) = 0$ for $|x| \geq 2$ and $0 \leq \chi(x) \leq 1$ for $1 \leq |x| \leq 2$.

2 Construction of an approximate solution – the NLW case

2.1 Inverting the linearized operator

Linearizing (1.1) around $\boldsymbol{W}$, $\boldsymbol{u} = \boldsymbol{W} + \boldsymbol{h}$, one obtains

$$\partial_t \boldsymbol{h} = J \circ \mathrm{D}^2 E(\boldsymbol{W})\boldsymbol{h} = \begin{pmatrix} 0 & \mathrm{Id} \\ -L & 0 \end{pmatrix} \boldsymbol{h},$$

where L is the Schrödinger operator

$$Lh := (-\Delta - f'(W))h = (-\Delta - 2W)h.$$

We introduce the following notation for the generators of the $\dot{H}^1$-critical and the L^2-critical scale change:

$$\Lambda := 2 + x \cdot \nabla, \qquad \Lambda_0 := 3 + x \cdot \nabla.$$

This is coherent with the definition of ΛW. Notice that $L(\Lambda W) = \frac{\mathrm{d}}{\mathrm{d}\lambda}\big|_{\lambda=1}\big(-\Delta W_\lambda - f(W_\lambda)\big) = 0$.

We fix $\mathcal{Z} \in C_0^\infty$ such that

$$\langle \mathcal{Z}, \Lambda W\rangle > 0, \qquad \langle \mathcal{Z}, \mathcal{Y}\rangle = 0.$$

We will use this function to define appropriate orthogonality conditions.

We denote also

$$\kappa := \Big(-\frac{\langle \Lambda W, f'(W)\rangle}{\langle \Lambda W, \Lambda W\rangle}\Big)^{\frac{1}{2}} = \sqrt{\frac{5}{4}}. \tag{2.1}$$

Lemma 2.1. *There exist radial rational functions $P(x), Q(x) \in C^\infty(\mathbb{R}^6)$ such that*

$$LP = \kappa^2 \Lambda W + f'(W), \qquad LQ = -\Lambda_0 \Lambda W, \tag{2.2}$$

$$\langle \mathcal{Z}, P\rangle = \langle \mathcal{Z}, Q\rangle = 0, \tag{2.3}$$

$$P(x) \sim |x|^{-2}, \qquad Q(x) \sim |x|^{-2} \qquad \textit{as } |x| \to +\infty. \tag{2.4}$$

Proof. By a direct computation one checks that the functions

$$\widetilde{P}(x) := \big(1 + \frac{|x|^2}{24}\big)^{-3} \cdot \big(1 - 10 \cdot \frac{|x|^2}{24} - 3 \cdot \big(\frac{|x|^2}{24}\big)^2\big),$$

$$\widetilde{Q}(x) := \big(1 + \frac{|x|^2}{24}\big)^{-3} \cdot \big(1 + 11 \cdot \frac{|x|^2}{24} - 12 \cdot \big(\frac{|x|^2}{24}\big)^2\big)$$

satisfy (2.2). Adding suitable multiples of ΛW to both functions we obtain P and Q satisfying (2.3). The formulas defining $\widetilde{P}$ and $\widetilde{Q}$ directly imply (2.4). □

Remark 2.2. Note that (2.4) is closely related to the Fredholm conditions $\langle \Lambda W, \kappa^2 \Lambda W + f'(W)\rangle = 0$ and $\langle \Lambda W, -\Lambda_0 \Lambda W\rangle = 0$, see Lemma 5.1 or [40, Proposition 2.1] for a more systematic presentation.

2.2 Formal computation

The usual method of performing a formal analysis of blow-up solutions is to search a series expansion with respect to a small scalar parameter depending on time and converging to 0 at blow-up. In our case the blow-up time is $-\infty$. If $u(t) \simeq W + W_{\lambda(t)}$, then $\partial_t u(t) \simeq -\lambda'(t)\Lambda W_{\underline{\lambda(t)}}$, hence

$$\boldsymbol{u}(t) \simeq (W + W_{\lambda(t)}, 0) - \lambda'(t) \cdot (0, \Lambda W_{\underline{\lambda(t)}}) = \boldsymbol{W} + \boldsymbol{U}^{(0)}_{\lambda(t)} + b(t) \cdot \boldsymbol{U}^{(1)}_{\lambda(t)},$$

with $b(t) := \lambda'(t)$, $\boldsymbol{U}^{(0)} := (W, 0)$ and $\boldsymbol{U}^{(1)} := (0, -\Lambda W)$. This suggests considering $b(t) = \lambda'(t)$ as the small parameter with respect to which the formal expansion should be sought. Hence, we make the ansatz

$$\boldsymbol{u}(t) = \boldsymbol{W} + \boldsymbol{U}^{(0)}_{\lambda(t)} + b(t) \cdot \boldsymbol{U}^{(1)}_{\lambda(t)} + b(t)^2 \cdot \boldsymbol{U}^{(2)}_{\lambda(t)},$$

and try to find the conditions under which a satisfactory candidate for $\boldsymbol{U}^{(2)} = (U^{(2)}, \dot{U}^{(2)})$ can be proposed. Neglecting irrelevant terms and replacing $\lambda'(t)$ by $b(t)$, we compute

$$\partial_t^2 u(t) = -b'(t)(\Lambda W)_{\underline{\lambda(t)}} + \frac{b(t)^2}{\lambda(t)}(\Lambda_0 \Lambda W)_{\underline{\lambda(t)}} + \text{lot}.$$

On the other hand, using the fact that $f(W+W_\lambda) = f(W)+f(W_\lambda)+f'(W_\lambda)W \simeq f(W)+f(W_\lambda)+f'(W_\lambda)$ for $\lambda \ll 1$ and $f'(W_\lambda) = \lambda f'(W)_{\underline{\lambda}}$, we get

$$\Delta u(t) + f(u(t)) = -\frac{b(t)^2}{\lambda(t)}(LU^{(2)})_{\underline{\lambda(t)}} + \lambda(t)f'(W)_{\underline{\lambda(t)}} + \mathrm{lot}.$$

We discover that, formally at least, we should have

$$LU^{(2)} = -\Lambda_0\Lambda W + \frac{\lambda(t)}{b(t)^2}\big(b'(t)\cdot\Lambda W + \lambda(t)\cdot f'(W)\big). \tag{2.5}$$

Lemma 2.1 shows that if $b'(t) = \kappa^2\lambda(t)$, then equation (2.5) has a decaying regular solution $U^{(2)} = Q + \frac{\lambda(t)^2}{b(t)^2}P$. The *formal parameter equations*

$$\lambda'(t) = b(t), \qquad b'(t) = \kappa^2\lambda(t)$$

have a solution

$$(\lambda_{\mathrm{app}}(t), b_{\mathrm{app}}(t)) = \big(\frac{1}{\kappa}\mathrm{e}^{-\kappa|t|}, \mathrm{e}^{-\kappa|t|}\big), \qquad t \le T_0 < 0.$$

In any space dimension N, ignoring the problems related to slow decay of W, a similar analysis would yield $b'(t) = \kappa^2\lambda(t)^{\frac{N-2}{4}}$. For $N < 6$ this leads to a finite time blow-up, which was studied in [40] for $N = 5$. For $N > 6$, we obtain a global solution $\lambda(t) \sim |t|^{-\frac{4}{N-6}}$, see Remark 1.4.

2.3 Bounds on the error of the ansatz

Let $I = [T, T_0]$ be a time interval, with $T \le T_0 < 0$ and $|T_0|$ large. Let $\lambda(t)$ and $\mu(t)$ be C^1 functions on $[T, T_0]$ such that

$$\lambda(T) = \frac{1}{\kappa}\mathrm{e}^{-\kappa|T|}, \quad \mu(T) = 1, \tag{2.6}$$

$$\frac{8}{9\kappa}\mathrm{e}^{-\kappa|t|} \le \lambda(t) \le \frac{9}{8\kappa}\mathrm{e}^{-\kappa|t|}, \qquad \frac{8}{9} \le \mu(t) \le \frac{9}{8}. \tag{2.7}$$

We define the *approximate solution* $\boldsymbol{\varphi}(t) = (\varphi(t), \dot\varphi(t)) : [T, T_0] \to \mathcal{E}$ by the formula

$$\begin{aligned}\varphi(t) &:= W_{\mu(t)} + W_{\lambda(t)} + S(t),\\ \dot\varphi(t) &:= -b(t)\Lambda W_{\underline{\lambda(t)}},\end{aligned}$$

where

$$b(t) := \mathrm{e}^{-\kappa|T|} + \kappa^2\int_T^t \frac{\lambda(\tau)}{\mu(\tau)^2}\,\mathrm{d}\tau, \qquad \text{for } t \in [T, T_0], \tag{2.8}$$

$$S(t) := \chi\cdot\Big(\frac{\lambda(t)^2}{\mu(t)^2}P_{\lambda(t)} + b(t)^2Q_{\lambda(t)}\Big), \qquad \text{for } t \in [T, T_0].$$

From (2.7) we get $\left(\frac{8}{9}\right)^3\frac{1}{\kappa}\mathrm{e}^{-\kappa|t|} \le \frac{\lambda(t)}{\mu(t)^2} \le \left(\frac{9}{8}\right)^3\frac{1}{\kappa}\mathrm{e}^{-\kappa|t|}$. Integrating we get the following bound for $b(t)$, $t \in [T, T_0]$:

$$\begin{aligned}\Big(\frac{8}{9}\Big)^3\mathrm{e}^{-\kappa|t|} &< \mathrm{e}^{-\kappa|T|} + \Big(\frac{8}{9}\Big)^3(\mathrm{e}^{-\kappa|t|} - \mathrm{e}^{-\kappa|T|})\\ &\le b(t) \le \mathrm{e}^{-\kappa|T|} + \Big(\frac{9}{8}\Big)^3(\mathrm{e}^{-\kappa|t|} - \mathrm{e}^{-\kappa|T|}) < \Big(\frac{9}{8}\Big)^3\mathrm{e}^{-\kappa|t|}.\end{aligned} \tag{2.9}$$

From (2.4) we obtain

$$
\begin{aligned}
\|\chi\cdot P_{\underline{\lambda}}\|_{\dot H^1} &\simeq \|\partial_r(\chi\cdot P_{\underline{\lambda}})\|_{L^2(r^5\,\mathrm{d}r)} \lesssim \|P_{\underline{\lambda}}\|_{L^2(0\le r\le 2)} + \|\frac{1}{\lambda}(\partial_r P)_{\underline{\lambda}}\|_{L^2(0\le r\le 2)} \\
&= \|P\|_{L^5(0\le r\le \frac{2}{\lambda})} + \frac{1}{\lambda}\|\partial_r P\|_{L^2(0\le r\le \frac{2}{\lambda})} \\
&\lesssim \Big(\int_0^{2/\lambda}(1+r^2)^{-2}r^5\,\mathrm{d}r\Big)^{\frac12} + \frac{1}{\lambda}\cdot\Big(\int_0^{2/\lambda}(1+r^3)^{-2}r^5\,\mathrm{d}r\Big)^{\frac12} \lesssim \frac{1}{\lambda}|\log\lambda|^{\frac12},
\end{aligned}
\tag{2.10}
$$

and analogously $\|\chi\cdot Q_{\underline{\lambda}}\|_{\dot H^1} \lesssim \frac{1}{\lambda}|\log\lambda|^{\frac12}$, hence

$$
\|S(t)\|_{\dot H^1} \le \frac{\lambda^3}{\mu^2}\|\chi P_{\underline{\lambda}}\|_{\dot H^1} + \lambda b^2\|\chi Q_{\underline{\lambda}}\|_{\dot H^1} \lesssim \mathrm{e}^{-3\kappa|t|}\cdot\frac{1}{\lambda}|\log\lambda|^{\frac12} \lesssim \sqrt{|t|}\cdot\mathrm{e}^{-2\kappa|t|}.
$$

Thus for any $c>0$ there exists T_0 such that if $T<T_0$ then

$$
\|S(t)\|_{\dot H^1} \le c\cdot \mathrm{e}^{-\frac32\kappa|t|}, \qquad \text{for } t\in[T,T_0]. \tag{2.11}
$$

Note also that $\|P_\lambda\|_{L^\infty} + \|Q_\lambda\|_{L^\infty} \lesssim \lambda^{-2}$, hence $S(t)$ is bounded in L^∞.

Since $\mathcal{Z}$ has compact support, for sufficiently small λ, (2.3) implies

$$
\langle \mathcal{Z}_{\underline{\lambda(t)}}, S(t)\rangle = 0. \tag{2.12}
$$

We denote

$$
\begin{aligned}
\boldsymbol{\psi}(t) = (\psi(t),\dot\psi(t)) &:= \partial_t\boldsymbol{\varphi}(t) - \mathrm{D}E(\boldsymbol{\varphi}(t)) \\
&= \big(\partial_t\varphi(t) - \dot\varphi(t), \partial_t\dot\varphi(t) - (\Delta\varphi(t) + f(\varphi(t)))\big).
\end{aligned}
\tag{2.13}
$$

This function describes how much $\boldsymbol{\varphi}(t)$ fails to be an exact solution of (1.1). Before we prove bounds on $\boldsymbol{\psi}(t)$, we gather in the next elementary lemma pointwise inequalities used in various places in the text.

Lemma 2.3. *Let $k,l,m\in\mathbb{R}$. Then*

$$
|f'(k+l) - f'(k)| \le f'(l), \tag{2.14}
$$

$$
|f(k+l) - f(k) - f'(k)l| \le 5|f(l)|, \tag{2.15}
$$

$$
|F(k+l) - F(k) - f(k)l - \frac12 f'(k)l^2| \le 5F(l). \tag{2.16}
$$

Proof. Inequality (2.14) is well-known. Bounds (2.15) holds for $k=0$, hence (by homogeneity) we may assume that $k=1$. For $|l|\le 1$ we have $|f(1+l)-f(1)-f'(1)l| = |(1+l)^2-1-2l| = 2l^2 \le 5|f(l)|$ and for $|l|\ge 1$ we find $|f(1+l)-f(1)-f'(1)l| \le (1+l)^2+1+2|l| \le 5|f(l)|$. Bound (2.16) follows by integrating (2.15). □

Lemma 2.4. *Suppose that for $t\in[T,T_0]$ there holds $|\lambda'(t)|\lesssim \mathrm{e}^{-\kappa|t|}$ and $|\mu'(t)|\lesssim\mathrm{e}^{-\kappa|t|}$. Then*

$$
\|\psi(t) + \mu'(t)\frac{1}{\mu(t)}\Lambda W_{\mu(t)} + (\lambda'(t)-b(t))\frac{1}{\lambda(t)}\Lambda W_{\lambda(t)}\|_{\dot H^1} \lesssim \mathrm{e}^{-\frac32\kappa|t|}, \tag{2.17}
$$

$$
\|\dot\psi(t) - \frac{b(t)}{\lambda(t)}(\lambda'(t)-b(t))\Lambda_0\Lambda W_{\underline{\lambda(t)}}\|_{L^2} \lesssim \mathrm{e}^{-\frac32\kappa|t|}, \tag{2.18}
$$

$$
\|(-\Delta - f'(\varphi(t)))\psi(t)\|_{\dot H^{-1}} \lesssim \mathrm{e}^{-\frac32\kappa|t|}. \tag{2.19}
$$

Proof. Using the definitions of $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$ we find

$$\begin{aligned}
&\psi + \mu'\Lambda W_{\underline{\mu}} + (\lambda' - b)\Lambda W_{\underline{\lambda}} = \partial_t\varphi - \dot\varphi + \mu'\Lambda W_{\underline{\mu}} + (\lambda' - b)\Lambda W_{\underline{\lambda}} \\
&= -\mu'\Lambda W_{\underline{\mu}} - \lambda'\Lambda W_{\underline{\lambda}} + \partial_t S + b\Lambda W_{\underline{\lambda}} + (\lambda' - b)\Lambda W_{\underline{\lambda}} \\
&= \chi\cdot\big(-2\mu'\frac{\lambda^3}{\mu^3}P_{\underline{\lambda}} + 2\lambda'\frac{\lambda^2}{\mu^2}P_{\underline{\lambda}} - \lambda'\frac{\lambda^2}{\mu^2}\Lambda P_{\underline{\lambda}} + 2b'b\lambda Q_{\underline{\lambda}} - \lambda' b^2\Lambda Q_{\underline{\lambda}}\big).
\end{aligned}$$

Since ΛP and ΛQ are rational functions decaying like r^{-2}, we have $\|\chi\cdot\Lambda P_{\underline{\lambda}}\|_{\dot H^1} \lesssim \sqrt{|t|}\cdot e^{\kappa|t|}$ and $\|\chi\cdot\Lambda Q_{\underline{\lambda}}\|_{\dot H^1} \lesssim \sqrt{|t|}\cdot e^{\kappa|t|}$, see (2.10). This implies (2.17) because $|\lambda|, |b|, |\lambda'|, |b'|, |\mu'| \lesssim e^{-\kappa|t|}$.

In order to prove (2.18), we consider separately the regions $|x| \le \sqrt{\lambda}$ and $|x| \ge \sqrt{\lambda}$. The first step is to treat the nonlinearity, that is to show that

$$\begin{aligned}
&\|f(\varphi) - f(W_\mu) - f(W_\lambda) - f'(W_\lambda)W_\mu \\
&\quad - \frac{\lambda^2}{\mu^2}f'(W_\lambda)P_\lambda - b^2 f'(W_\lambda)Q_\lambda\|_{L^2(|x|\le\sqrt{\lambda})} \lesssim e^{-\frac{3}{2}\kappa|t|}.
\end{aligned} \tag{2.20}$$

Applying (2.15) with $k = W_\lambda$ and $l = W_\mu + S$ we get

$$\big|f(\varphi) - f(W_\lambda) - f'(W_\lambda)(W_\mu + \lambda^2\mu^{-2}\cdot P_\lambda + b^2\cdot Q_\lambda)\big| \lesssim |f(W_\mu)| + |f(\lambda^2\mu^{-2}P_\lambda + b^2 Q_\lambda)|,$$

which is bounded in L^∞, hence bounded by $\lambda^{\frac{3}{2}} \sim e^{-\frac{3}{2}\kappa|t|}$ in $L^2(|x| \le \sqrt{\lambda})$. This proves (2.20). Now we check that

$$\big\|f'(W_\lambda)W_\mu - \frac{1}{\mu^2}f'(W_\lambda)\big\|_{L^2} \lesssim e^{-\frac{3}{2}\kappa|t|}. \tag{2.21}$$

Indeed, for $|x| \le \sqrt{\lambda}$ we have $|W_\mu(x) - \frac{1}{\mu^2}| = |W_\mu(x) - W_\mu(0)| \lesssim |x|^2 \le \lambda \sim e^{-\kappa|t|}$, hence

$$\big\|f'(W_\lambda)W_\mu - \frac{1}{\mu^2}f'(W_\lambda)\big\|_{L^2(|x|\le\sqrt{\lambda})} \lesssim \big\|W_\mu - \frac{1}{\mu^2}\big\|_{L^\infty(|x|\le\sqrt{\lambda})}\cdot\|W_\lambda\|_{L^2} \lesssim e^{-\kappa|t|}\cdot\lambda \ll e^{-\frac{3}{2}\kappa|t|}.$$

From (2.20) and (2.21) we obtain

$$\|f(\varphi) - f(W_\mu) - f(W_\lambda) - \mu^{-2}f'(W_\lambda) - \lambda^2\mu^{-2}f'(W_\lambda)P_\lambda - b^2 f'(W_\lambda)Q_\lambda\|_{L^2} \lesssim e^{-\frac{3}{2}\kappa|t|}. \tag{2.22}$$

Since $\chi = 1$ in the region $|x| \le \sqrt{\lambda}$, we have $\Delta\varphi = \Delta(W_\mu) + \Delta(W_\lambda) + \lambda^2\mu^{-2}\Delta(P_\lambda) + b^2\Delta(Q_\lambda)$. From this and (2.22), using the fact that $\Delta(W_\mu) + f(W_\mu) = \Delta(W_\lambda) + f(W_\lambda) = 0$, we get

$$\begin{aligned}
&\big\|\Delta\varphi + f(\varphi) - \mu^{-2}\big(\lambda^2\Delta(P_\lambda) + \lambda^2 f'(W_\lambda)P_\lambda + f'(W_\lambda)\big) \\
&\quad - \big(b^2\Delta(Q_\lambda) + b^2 f'(W_\lambda)Q_\lambda\big)\big\|_{L^2(|x|\le\sqrt{\lambda})} \lesssim e^{-\frac{3}{2}\kappa|t|}.
\end{aligned} \tag{2.23}$$

But formula (2.2) gives

$$\begin{aligned}
\lambda^2\Delta(P_\lambda) + \lambda^2 f'(W_\lambda)P_\lambda &= (-LP)_\lambda = -\kappa^2\Lambda W_\lambda - f'(W_\lambda) = -\kappa^2\lambda\Lambda W_{\underline{\lambda}} - f'(W_\lambda), \\
b^2\Delta(Q_\lambda) + b^2 f'(W_\lambda)Q_\lambda &= \frac{b^2}{\lambda^2}(-LQ)_\lambda = \frac{b^2}{\lambda}\Lambda_0\Lambda W_{\underline{\lambda}},
\end{aligned}$$

hence we can rewrite (2.23) as

$$\big\|\Delta\varphi + f(\varphi) + \frac{\kappa^2\lambda}{\mu^2}\Lambda W_{\underline{\lambda}} - \frac{b^2}{\lambda}\Lambda_0\Lambda W_{\underline{\lambda}}\big\|_{L^2(|x|\le\sqrt{\lambda})} \lesssim e^{-\frac{3}{2}\kappa|t|}. \tag{2.24}$$

We have $\partial_t\dot\varphi = -b'\Lambda W_{\underline\lambda} + \frac{\lambda' b}{\lambda}\Lambda_0\Lambda W_{\underline\lambda}$, thus

$$\begin{aligned}-\dot\psi + \frac{b}{\lambda}(\lambda'-b)\Lambda_0\Lambda W_{\underline\lambda} &= \Delta\varphi + f(\varphi) + b'\Lambda W_{\underline\lambda} - \frac{b\lambda'}{\lambda}\Lambda_0\Lambda W_{\underline\lambda} + \frac{b}{\lambda}(\lambda'-b)\Lambda_0\Lambda W_{\underline\lambda} \\ &= \Delta\varphi + f(\varphi) + \frac{\kappa^2\lambda}{\mu^2}\Lambda W_{\underline\lambda} - \frac{b^2}{\lambda}\Lambda_0\Lambda W_{\underline\lambda},\end{aligned} \tag{2.25}$$

so (2.24) yields (2.18) in the region $|x|\leq\sqrt\lambda$.

Consider now the region $|x|\geq\sqrt\lambda$. First we show that

$$\|\Delta\varphi - \Delta(W_\mu)\|_{L^2(|x|\geq\sqrt\lambda)} \lesssim \mathrm{e}^{-\frac32\kappa|t|}. \tag{2.26}$$

To this end, we compute

$$\begin{aligned}\|\Delta(W_\lambda)\|_{L^2(|x|\geq\sqrt\lambda)} = \|f(W_\lambda)\|_{L^2(|x|\geq\sqrt\lambda)} &= \frac{1}{\lambda}\|f(W)\|_{L^2(|x|\geq 1/\sqrt\lambda)} \\ &\lesssim \frac{1}{\lambda}\Big(\int_{1/\sqrt\lambda}^{+\infty} r^{-16}r^5\,\mathrm{d}r\Big)^{\frac12} \lesssim \frac1\lambda\lambda^{\frac52} \lesssim \mathrm{e}^{-\frac32\kappa|t|}.\end{aligned}$$

We need to show that $\|\Delta(\chi\cdot P_\lambda)\|_{L^2(|x|\geq\sqrt\lambda)} \lesssim \mathrm{e}^{\frac12\kappa|t|}$ and $\|\Delta(\chi\cdot Q_\lambda)\|_{L^2(|x|\geq\sqrt\lambda)} \lesssim \mathrm{e}^{\frac12\kappa|t|}$. We will prove the first bound (the second is exactly the same). Notice that $|P_\lambda(x)| \lesssim \frac{1}{\lambda^2}\cdot\frac{\lambda^2}{|x|^2} = |x|^{-2}$ and similarly $|\nabla(P_\lambda)(x)| \lesssim |x|^{-3}$, $|\nabla^2(P_\lambda)(x)| \lesssim |x|^{-4}$, hence we have a pointwise bound

$$|\Delta(\chi P_\lambda)| \lesssim |\nabla^2\chi|\cdot|P_\lambda| + |\nabla\chi|\cdot|\nabla(P_\lambda)| + |\chi|\cdot|\nabla^2(P_\lambda)| \lesssim |\nabla^2\chi|\cdot|x|^{-2} + |\nabla\chi|\cdot|x|^{-3} + |\chi|\cdot|x|^{-4}.$$

Of course $\big\||\nabla^2\chi|\cdot|x|^{-2}\big\|_{L^2(|x|\geq\sqrt\lambda)} + \big\||\nabla\chi|\cdot|x|^{-3}\big\|_{L^2(|x|\geq\sqrt\lambda)} \lesssim 1 \ll \mathrm{e}^{\frac12\kappa|t|}$ and we are left with the last term. We compute

$$\big\||\chi|\cdot|x|^{-4}\|_{L^2(|x|\geq\sqrt\lambda)} \lesssim \Big(\int_{\sqrt\lambda}^{2} r^{-8}r^5\,\mathrm{d}r\Big)^{\frac12} \lesssim \lambda^{-\frac12} \lesssim \mathrm{e}^{\frac12\kappa|t|}.$$

This finishes the proof of (2.26).

Applying (2.15) with $k = W_\mu$ and $l = W_\lambda + S$ we get

$$|f(\varphi) - f(W_\mu)| \lesssim f'(W_\mu)\cdot(|W_\lambda| + |S|) + |f(W_\lambda)| + |f(S)| \lesssim |W_\lambda| + |S|,$$

where the last estimate follows from the fact that $\|W_\lambda\|_{L^\infty} + \|S\|_{L^\infty} \lesssim 1$ for $|x|\geq\sqrt\lambda$.

We have $\|\chi\cdot P_{\underline\lambda}\|_{L^2} \lesssim \Big(\int_0^{\frac2\lambda}(r^{-2})^2 r^5\,\mathrm{d}r\Big)^{\frac12} \sim \lambda^{-1}$, and similarly $\|\chi\cdot Q_{\underline\lambda}\|_{L^2} \lesssim \lambda^{-1}$, which implies $\|S\|_{L^2} \ll \mathrm{e}^{-\frac32\kappa|t|}$. There holds also

$$\|W_\lambda\|_{L^2(|x|\geq\sqrt\lambda)} = \lambda\|W\|_{L^2(|x|\geq 1/\sqrt\lambda)} \lesssim \lambda\Big(\int_{1/\sqrt\lambda}^{\infty} r^{-8}r^5\,\mathrm{d}r\Big)^{\frac12} \sim \lambda^{\frac32} \sim \mathrm{e}^{-\frac32\kappa|t|}, \tag{2.27}$$

hence $\|f(\varphi) - f(\mu)\|_{L^2(|x|\geq\sqrt\lambda)} \lesssim \mathrm{e}^{-\frac32\kappa|t|}$. Together with (2.26) this yields

$$\|\Delta\varphi + f(\varphi)\|_{L^2(|x|\geq\sqrt\lambda)} \lesssim \mathrm{e}^{-\frac32\kappa|t|}.$$

The same computation as in (2.27) gives

$$\|\Lambda W_{\underline\lambda}\|_{L^2(|x|\geq\sqrt\lambda)} + \|\Lambda_0\Lambda W_{\underline\lambda}\|_{L^2(|x|\geq\sqrt\lambda)} \lesssim \mathrm{e}^{-\frac12\kappa|t|},$$

hence (2.25) implies that (2.18) holds also in the region $|x|\geq\sqrt\lambda$.

We are left with (2.19). From (2.17) it follows that it suffices to check that

$$\Big\|(-\Delta - f'(\varphi))(\lambda' - b)\frac{1}{\lambda}\Lambda W_\lambda\Big\|_{\dot H^{-1}} \lesssim \mathrm{e}^{-\frac{3}{2}\kappa|t|} \tag{2.28}$$

and

$$\Big\|(-\Delta - f'(\varphi))\mu'\frac{1}{\mu}\Lambda W_\mu\Big\|_{\dot H^{-1}} \lesssim \mathrm{e}^{-\frac{3}{2}\kappa|t|}. \tag{2.29}$$

We start with (2.28). Since $|\lambda' - b| \lesssim \mathrm{e}^{-\kappa|t|} \lesssim \lambda$, we need to show that

$$\|(-\Delta - f'(\varphi))\Lambda W_\lambda\|_{\dot H^{-1}} \lesssim \mathrm{e}^{-\frac{3}{2}\kappa|t|}. \tag{2.30}$$

By Hölder inequality

$$\|f'(W_\mu)\Lambda W_\lambda\|_{L^{\frac{3}{2}}} \leq \|f'(W_\mu)\|_{L^{12}} \cdot \|\Lambda W_\lambda\|_{L^{\frac{12}{7}}} \lesssim \lambda^{\frac{3}{2}} \lesssim \mathrm{e}^{-\frac{3}{2}\kappa|t|}. \tag{2.31}$$

Since $|f'(W_\lambda + W_\mu) - f'(W_\lambda)| \leq f'(W_\mu)$, we obtain

$$\|(f'(W_\lambda + W_\mu) - f'(W_\lambda))\Lambda W_\lambda\|_{\dot H^{-1}} \lesssim \|(f'(W_\lambda + W_\mu) - f'(W_\lambda))\Lambda W_\lambda\|_{L^{\frac{3}{2}}} \lesssim \mathrm{e}^{-\frac{3}{2}\kappa|t|}.$$

As noted earlier $(-\Delta - f'(W_\lambda))\Lambda W_\lambda = 0$, hence

$$\|(-\Delta - f'(W_\lambda + W_\mu))\Lambda W_\lambda\|_{\dot H^{-1}} \lesssim \mathrm{e}^{-\frac{3}{2}\kappa|t|}.$$

From (2.11) we have $\|f'(\varphi) - f'(W_\lambda + W_\mu)\|_{L^3} \lesssim \mathrm{e}^{-\frac{3}{2}\kappa|t|}$. This implies (2.30).

The proof of (2.29) is similar. It suffices to check that $\|f'(W_\lambda)\Lambda W_\mu\|_{\dot H^{-1}} \lesssim \mathrm{e}^{-\frac{1}{2}\kappa|t|}$, and in fact we even have the bound $\lesssim \mathrm{e}^{-\frac{3}{2}\kappa|t|}$, with the same proof as in (2.31). □

3 Bootstrap control of the error term – the NLW case

In the preceding section we defined *approximate* solutions of (1.1). In the present section we consider *exact* solutions of (1.1), with some specific initial data prescribed at $t = T$, with $T \to -\infty$. Our goal is to control the evolution of this solution up to a time T_0 independent of T.

For technical reasons we will require the initial data to belong to the space $X^1 \times H^1$, where $X^1 := \dot H^2 \cap \dot H^1$. This regularity is preserved by the flow, see Proposition A.1.

3.1 Set-up of the bootstrap

It is known that $L = -\Delta - f'(W)$ has exactly one strictly negative simple eigenvalue which we denote $-\nu^2$ (we take $\nu > 0$). We denote the corresponding positive eigenfunction $\mathcal{Y}$, normalized so that $\|\mathcal{Y}\|_{L^2} = 1$. By elliptic regularity $\mathcal{Y}$ is smooth and by Agmon estimates it decays exponentially. Self-adjointness of L implies that

$$\langle \mathcal{Y}, \Lambda W\rangle = 0.$$

Note that

$$\nu < 1. \tag{3.1}$$

Indeed, it is well-known that $-\Delta - W \geq 0$, with a one-dimensional kernel generated by W. Since $1 - W(x) > 0$ almost everywhere, for any $h \neq 0$ we have

$$\langle h, Lh\rangle + \langle h, h\rangle = \langle(-\Delta - 2W + 1)h, h\rangle > \langle(-\Delta - W)h, h\rangle \geq 0.$$

We define

$$\mathcal{Y}^- := \big(\frac{1}{\nu}\mathcal{Y}, -\mathcal{Y}\big), \qquad \mathcal{Y}^+ := \big(\frac{1}{\nu}\mathcal{Y}, \mathcal{Y}\big),$$
$$\alpha^- := \frac{\nu}{2}J\mathcal{Y}^+ = \frac{1}{2}(\nu\mathcal{Y}, -\mathcal{Y}), \qquad \alpha^+ := -\frac{\nu}{2}J\mathcal{Y}^- = \frac{1}{2}(\nu\mathcal{Y}, \mathcal{Y}).$$

We have $J \circ \mathrm{D}^2 E(\boldsymbol{W}) = \begin{pmatrix} 0 & \mathrm{Id} \\ -L & 0 \end{pmatrix}$. A short computation shows that

$$J \circ \mathrm{D}^2 E(\boldsymbol{W})\mathcal{Y}^- = -\nu\mathcal{Y}^-, \qquad J \circ \mathrm{D}^2 E(\boldsymbol{W})\mathcal{Y}^+ = \nu\mathcal{Y}^+$$

and

$$\langle \alpha^-, J \circ \mathrm{D}^2 E(\boldsymbol{W})\boldsymbol{h}\rangle = -\nu\langle\alpha^-, \boldsymbol{h}\rangle, \qquad \langle \alpha^+, J \circ \mathrm{D}^2 E(\boldsymbol{W})\boldsymbol{h}\rangle = \nu\langle\alpha^+, \boldsymbol{h}\rangle, \qquad \forall \boldsymbol{h} \in \mathcal{E}.$$

We will think of α^- and α^+ as linear forms on $\mathcal{E}$. Notice that $\langle\alpha^-, \mathcal{Y}^-\rangle = \langle\alpha^+, \mathcal{Y}^+\rangle = 1$ and $\langle\alpha^-, \mathcal{Y}^+\rangle = \langle\alpha^+, \mathcal{Y}^-\rangle = 0$.

The rescaled versions of these objects are

$$\mathcal{Y}_\lambda^- := \big(\frac{1}{\nu}\mathcal{Y}_\lambda, -\mathcal{Y}_{\underline{\lambda}}\big), \qquad \mathcal{Y}_\lambda^+ := \big(\frac{1}{\nu}\mathcal{Y}_\lambda, \mathcal{Y}_{\underline{\lambda}}\big),$$
$$\alpha_\lambda^- := \frac{\nu}{2\lambda}J\mathcal{Y}_\lambda^+ = \frac{1}{2}\big(\frac{\nu}{\lambda}\mathcal{Y}_{\underline{\lambda}}, -\mathcal{Y}_{\underline{\lambda}}\big), \qquad \alpha_\lambda^+ := -\frac{\nu}{2\lambda}J\mathcal{Y}_\lambda^- = \frac{1}{2}\big(\frac{\nu}{\lambda}\mathcal{Y}_{\underline{\lambda}}, \mathcal{Y}_{\underline{\lambda}}\big). \tag{3.2}$$

The scaling is chosen so that $\langle\alpha_\lambda^-, \mathcal{Y}_\lambda^-\rangle = \langle\alpha_\lambda^+, \mathcal{Y}_\lambda^+\rangle = 1$. We have

$$J \circ \mathrm{D}^2 E(\boldsymbol{W}_\lambda)\mathcal{Y}_\lambda^- = -\frac{\nu}{\lambda}\mathcal{Y}_\lambda^-, \qquad J \circ \mathrm{D}^2 E(\boldsymbol{W}_\lambda)\mathcal{Y}_\lambda^+ = \frac{\nu}{\lambda}\mathcal{Y}_\lambda^+ \tag{3.3}$$

and

$$\langle \alpha_\lambda^-, J \circ \mathrm{D}^2 E(\boldsymbol{W}_\lambda)\boldsymbol{h}\rangle = -\frac{\nu}{\lambda}\langle\alpha_\lambda^-, \boldsymbol{h}\rangle, \qquad \langle \alpha_\lambda^+, J \circ \mathrm{D}^2 E(\boldsymbol{W}_\lambda)\boldsymbol{h}\rangle = \frac{\nu}{\lambda}\langle\alpha_\lambda^+, \boldsymbol{h}\rangle, \qquad \forall \boldsymbol{h} \in \mathcal{E}. \tag{3.4}$$

We will need the following simple lemma in order to properly choose the initial data.

Lemma 3.1. *There exist universal constants $\eta, C > 0$ such that if $0 < \lambda < \eta \cdot \mu$, then for all $a_0 \in \mathbb{R}$ there exists $\boldsymbol{h}_0 \in X^1 \times H^1$ satisfying the orthogonality conditions $\langle \mathcal{Z}_\mu, h_0\rangle = \langle \mathcal{Z}_\lambda, h_0\rangle = 0$ and such that $\langle\alpha_\mu^+, \boldsymbol{h}_0\rangle = 0$, $\langle\alpha_\mu^-, \boldsymbol{h}_0\rangle = 0$, $\langle\alpha_\lambda^+, \boldsymbol{h}_0\rangle = a_0$, $\langle\alpha_\lambda^-, \boldsymbol{h}_0\rangle = 0$, $\|\boldsymbol{h}_0\|_\mathcal{E} \leq C|a_0|$.*

Proof. We consider functions of the form:

$$\boldsymbol{h}_0 := a_2^+\mathcal{Y}_\mu^+ + a_2^-\mathcal{Y}_\mu^- + b_2\Lambda\boldsymbol{W}_\mu + a_1^+\mathcal{Y}_\lambda^+ + a_1^-\mathcal{Y}_\lambda^- + b_1\Lambda\boldsymbol{W}_\lambda, \qquad a_2^+, a_2^-, b_2, a_1^+, a_1^-, b_1 \in \mathbb{R}.$$

Consider the linear map $\Phi : \mathbb{R}^6 \to \mathbb{R}^6$ defined as follows:

$$\Phi(a_2^+, a_2^-, b_2, a_1^+, a_1^-, b) := \big(\langle\alpha_\mu^+, \boldsymbol{h}_0\rangle, \langle\alpha_\mu^-, \boldsymbol{h}_0\rangle, \langle\frac{1}{\mu}\mathcal{Z}_{\underline{\mu}}, h_0\rangle, \langle\alpha_\lambda^+, \boldsymbol{h}_0\rangle, \langle\alpha_\lambda^-, \boldsymbol{h}_0\rangle, \langle\frac{1}{\lambda}\mathcal{Z}_{\underline{\lambda}}, h_0\rangle\big)$$

It is easy to check that the matrix of Φ is strictly diagonally dominant if η is small enough. □

We consider the solution $\boldsymbol{u}(t) = \boldsymbol{u}(a_0; t) : [T, T_+) \to \mathcal{E}$ of (1.1) with the initial data

$$\boldsymbol{u}(T) = \big(W_{\frac{1}{\kappa}\mathrm{e}^{-\kappa|T|}} + W + h_0, -\Lambda W_{\underline{\frac{1}{\kappa}\mathrm{e}^{-\kappa|T|}}}\big), \tag{3.5}$$

where h_0 is the function given by Lemma 3.1 with $\lambda = \frac{1}{\kappa}\mathrm{e}^{-\kappa|T|}$, $\mu = 1$ and some a_0 chosen later, satisfying

$$|a_0| \leq \mathrm{e}^{-\frac{3}{2}\kappa|T|}.$$

Note that the initial data depends continuously on a_0.

For $t \geq T$ we define the functions $\lambda(t)$ and $\mu(t)$ as the solutions of the following system of ordinary differential equations with the initial data $\mu(T) = 1$ and $\lambda(T) = \frac{1}{\kappa}\mathrm{e}^{-\kappa|T|}$:

$$\begin{pmatrix} \langle \mathcal{Z}_{\underline{\lambda}}, \Lambda W_{\underline{\lambda}}\rangle - \langle \frac{1}{\lambda}\Lambda_0\mathcal{Z}_{\underline{\lambda}}, h\rangle & \langle \mathcal{Z}_{\underline{\lambda}}, \Lambda W_{\underline{\mu}}\rangle \\ \langle \mathcal{Z}_{\underline{\mu}}, \Lambda W_{\underline{\lambda}}\rangle & \langle \mathcal{Z}_{\underline{\mu}}, \Lambda W_{\underline{\mu}}\rangle - \langle \frac{1}{\mu}\Lambda_0\mathcal{Z}_{\underline{\mu}}, h\rangle \end{pmatrix} \cdot \begin{pmatrix} \lambda' \\ \mu' \end{pmatrix} = \begin{pmatrix} -\langle \mathcal{Z}_{\underline{\lambda}}, \partial_t u\rangle \\ -\langle \mathcal{Z}_{\underline{\mu}}, \partial_t u\rangle \end{pmatrix}, \tag{3.6}$$

where

$$h = h(t) := \begin{cases} u(t) - W_{\mu(t)} - W_{\lambda(t)}, & \text{if } \|u(t) - W_{\mu(t)} - W_{\lambda(t)}\|_{\dot H^1} \leq \eta, \\ \frac{\eta}{\|u(t)-W_{\mu(t)}-W_{\lambda(t)}\|_{\dot H^1}}\big(u(t) - W_{\mu(t)} - W_{\lambda(t)}\big), & \text{if } \|u(t) - W_{\mu(t)} - W_{\lambda(t)}\|_{\dot H^1} \geq \eta \end{cases}$$

with a small constant $\eta > 0$. Notice that $\langle \mathcal{Z}_{\underline{\lambda}}, \Lambda W_{\underline{\lambda}}\rangle = \langle \mathcal{Z}_{\underline{\mu}}, \Lambda W_{\underline{\mu}}\rangle = \langle \mathcal{Z}, \Lambda W\rangle > 0$, $|\langle \frac{1}{\lambda}\Lambda_0\mathcal{Z}_{\underline{\lambda}}, h\rangle| + |\langle \frac{1}{\mu}\Lambda_0\mathcal{Z}_{\underline{\mu}}, h\rangle| \lesssim \|h\|_{\dot H^1}$ and $|\langle \mathcal{Z}_{\underline{\lambda}}, \Lambda W_{\underline{\mu}}\rangle| + |\langle \mathcal{Z}_{\underline{\mu}}, \Lambda W_{\underline{\lambda}}\rangle| \lesssim \lambda/\mu$. For $t < T_0$ bounds (2.7) imply that λ/μ is small, hence equation (3.6) defines a unique solution as long as (2.7) holds.

Remark 3.2. Actually the second case in the definition of $h(t)$ will never occur in our analysis, since the bootstrap assumptions imply that $\|h(t)\|_{\dot H^1}$ is small.

Suppose that $\lambda(t)$ and $\mu(t)$ are well defined and satisfy (2.7) for $t \in [T, T_1]$, where $T < T_1 \leq T_0$. Suppose also that $\|h(t)\|_{\dot H^1} < \eta$ for $t \in [T, T_1]$, which implies that $h(t) = u(t)$. Using (3.6) we find $\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathcal{Z}_{\underline{\mu(t)}}, h(t)\rangle = 0$ and $\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathcal{Z}_{\underline{\lambda(t)}}, h(t)\rangle = 0$. Since $\langle \mathcal{Z}_{\underline{\mu(T)}}, h(T)\rangle = 0$ and $\langle \mathcal{Z}_{\underline{\lambda(T)}}, h(T)\rangle = 0$, we obtain

$$\langle \mathcal{Z}_{\underline{\mu(t)}}, h(t)\rangle = \langle \mathcal{Z}_{\underline{\lambda(t)}}, h(t)\rangle = 0, \qquad \text{for } t \in [T, T_1]. \tag{3.7}$$

We denote $\boldsymbol{h}(t) := (h(t), \partial_t u(t))$, so that

$$\begin{cases} \partial_t h = \dot h + \mu'\Lambda W_{\underline{\mu}} + \lambda'\Lambda W_{\underline{\lambda}}, \\ \partial_t \dot h = \Delta h + f(W_\mu + W_\lambda + h) - f(W_\mu) - f(W_\lambda). \end{cases} \tag{3.8}$$

We define the function $b(t) : [T, T_1] \to \mathbb{R}$ by formula (2.8) and decompose

$$\boldsymbol{u}(t) = \boldsymbol{\varphi}(t) + \boldsymbol{g}(t), \qquad t \in [T, T_1].$$

By the definitions of $\boldsymbol{g}(t)$ and $\boldsymbol{\psi}(t)$, $\boldsymbol{g}(t)$ satisfies the differential equation

$$\partial_t \boldsymbol{g}(t) = J \circ \mathrm{D}E(\boldsymbol{\varphi}(t) + \boldsymbol{g}(t)) - J \circ \mathrm{D}E(\boldsymbol{\varphi}(t)) - \boldsymbol{\psi}(t). \tag{3.9}$$

Finally, we denote

$$\begin{aligned} a_1^+(t) &:= \langle \alpha^+_{\lambda(t)}, \boldsymbol{g}(t)\rangle, & a_1^-(t) &:= \langle \alpha^-_{\lambda(t)}, \boldsymbol{g}(t)\rangle, \\ a_2^+(t) &:= \langle \alpha^+_{\mu(t)}, \boldsymbol{g}(t)\rangle, & a_2^-(t) &:= \langle \alpha^-_{\mu(t)}, \boldsymbol{g}(t)\rangle. \end{aligned}$$

The rest of this section is devoted to the proof of the following bootstrap estimate, which is the heart of the whole construction.

Proposition 3.3. *There exist constants $C_0 > 0$ and $T_0 < 0$ (C_0 and $|T_0|$ large) with the following property. Let $T < T_1 < T_0$ and suppose that $\boldsymbol{u}(t) = \boldsymbol{\varphi}(t)+\boldsymbol{g}(t) \in C([T,T_1];X^1\times H^1)$ is a solution of* (1.1) *with initial data* (3.5) *such that for $t \in [T,T_1]$ condition* (2.7) *is satisfied and*

$$\|\boldsymbol{g}(t)\|_{\mathcal{E}} \leq C_0 \cdot \mathrm{e}^{-\frac{3}{2}\kappa|t|}, \tag{3.10}$$

$$|a_1^+(t)| \leq \mathrm{e}^{-\frac{3}{2}\kappa|t|}. \tag{3.11}$$

Then for $t \in [T,T_1]$ there holds

$$\|\boldsymbol{g}(t)\|_{\mathcal{E}} \leq \frac{1}{2}C_0\mathrm{e}^{-\frac{3}{2}\kappa|t|}, \tag{3.12}$$

$$\big|\lambda(t) - \frac{1}{\kappa}\mathrm{e}^{-\kappa|t|}\big| + |\mu(t)-1| \lesssim C_0\mathrm{e}^{-\frac{3}{2}\kappa|t|}. \tag{3.13}$$

Remark 3.4. Notice that (3.12) and (3.13) are strictly stronger than (3.10) and (2.7) respectively, which will be crucial for closing the bootstrap in Subsection 3.6.

Remark 3.5. The same conclusion should be true without the assumption of $X^1 \times H^1$ regularity, by means of a standard approximation procedure (both the assumptions and the conclusion are continuous for the topology $\|\cdot\|_{\mathcal{E}}$).

3.2 Modulation

Lemma 3.6. *Under assumptions* (2.7) *and* (3.10), *for $t \in [T,T_1]$ there holds*

$$\big\langle \frac{1}{\lambda(t)}\mathcal{Z}_{\underline{\lambda(t)}}, g(t)\big\rangle = 0, \quad \big|\big\langle \frac{1}{\mu(t)}\mathcal{Z}_{\underline{\mu(t)}}, g(t)\big\rangle\big| \lesssim c\cdot \mathrm{e}^{-\frac{3}{2}\kappa|t|}, \tag{3.14}$$

$$|\lambda'(t) - b(t)| + |\mu'(t)| \lesssim \|\boldsymbol{g}(t)\|_{\mathcal{E}} + c\cdot\mathrm{e}^{-\frac{3}{2}\kappa|t|}, \tag{3.15}$$

with a constant c arbitrarily small.

Proof. We have $g(t) = h(t) - S(t)$. Since $\mathcal{Z}$ has compact support, (2.12) and (3.7) yield $\langle \mathcal{Z}_{\underline{\lambda}}, g\rangle = 0$. From $\|\frac{1}{\mu}\mathcal{Z}_{\underline{\mu}}\|_{L^\infty} \lesssim 1$ and $\|\chi\cdot\lambda^2 P_\lambda\|_{L^1} \leq \|\lambda^2 P_\lambda\|_{L^1(|x|\leq 2)} \lesssim \lambda^2 \sim \mathrm{e}^{-2\kappa|t|}$ (analogously $\|\chi\cdot b^2 Q_\lambda\|_{L^1} \lesssim \mathrm{e}^{-2\kappa|t|}$) we obtain $|\langle\frac{1}{\mu}\mathcal{Z}_{\underline{\mu}}, g\rangle| \lesssim \mathrm{e}^{-2\kappa|t|}$.

From (2.11) we have

$$\|g(t)-h(t)\|_{\dot{H}^1} \leq c\cdot \mathrm{e}^{-\frac{3}{2}\kappa|t|}, \qquad \text{with a small constant } c, \tag{3.16}$$

Using this, (3.6) yields

$$\begin{pmatrix} \langle\mathcal{Z}_{\underline{\lambda}}, \Lambda W_{\underline{\lambda}}\rangle - \langle\frac{1}{\lambda}\Lambda_0\mathcal{Z}_{\underline{\lambda}}, g\rangle & \langle\mathcal{Z}_{\underline{\lambda}}, \Lambda W_{\underline{\mu}}\rangle \\ \langle\mathcal{Z}_{\underline{\mu}}, \Lambda W_{\underline{\lambda}}\rangle & \langle\mathcal{Z}_{\underline{\mu}}, \Lambda W_{\underline{\mu}}\rangle - \langle\frac{1}{\mu}\Lambda_0\mathcal{Z}_{\underline{\mu}}, g\rangle \end{pmatrix}\cdot\begin{pmatrix}\lambda' - b \\ \mu'\end{pmatrix}$$
$$= \begin{pmatrix} -\langle\mathcal{Z}_{\underline{\lambda}}, \partial_t u + b\Lambda W_{\underline{\lambda}}\rangle \\ -\langle\mathcal{Z}_{\underline{\mu}}, \partial_t u + b\Lambda W_{\underline{\lambda}}\rangle\end{pmatrix} + o_{\mathbb{R}^2}(\mathrm{e}^{-\frac{3}{2}\kappa|t|}).$$

Since by definition $\partial_t u + b\Lambda W_{\underline{\lambda}} = \dot{g}$, inverting the matrix we get (3.15). □

Lemma 3.7. *Under assumptions* (2.7) *and* (3.10), *for $t \in [T,T_1]$ the functions $a_1^\pm(t)$ and $a_2^\pm(t)$ satisfy*

$$\big|\frac{\mathrm{d}}{\mathrm{d}t}a_1^+(t) - \frac{\nu}{\lambda(t)}a_1^+(t)\big| \leq c\cdot\mathrm{e}^{-\frac{1}{2}\kappa|t|}, \qquad \textit{with a small constant } c \tag{3.17}$$

$$|a_1^-(t)| \leq \mathrm{e}^{-\frac{3}{2}\kappa|t|}, \tag{3.18}$$

$$|a_2^\pm(t)| \leq \mathrm{e}^{-\frac{3}{2}\kappa|t|}. \tag{3.19}$$

Proof. Using the definition of $a_1^+(t)$ we compute

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}a_1^+(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\langle \alpha_{\lambda(t)}^+, \boldsymbol{g}(t)\rangle \\ &= \big\langle -\frac{\lambda'}{\lambda}(\Lambda_{\mathcal{E}^*}\alpha^+)_\lambda, \boldsymbol{g}\big\rangle + \langle \alpha_\lambda^+, J\circ \mathrm{D}E(\boldsymbol{\varphi}+\boldsymbol{g}) - J\circ \mathrm{D}E(\boldsymbol{\varphi}) - \boldsymbol{\psi}\rangle,\end{aligned} \tag{3.20}$$

where $\Lambda_{\mathcal{E}^*}\alpha^+ := -\frac{\partial}{\partial\lambda}\alpha_\lambda^+|_{\lambda=1}$. We have

$$|\langle(\Lambda_{\mathcal{E}^*}\alpha^+)_\lambda, \boldsymbol{g}\rangle| \lesssim \|\boldsymbol{g}\| \ll \mathrm{e}^{-\frac{1}{2}\kappa|t|}. \tag{3.21}$$

Since $\langle \alpha_\lambda^+, (\Lambda W_\lambda, 0)\rangle = 0$, using (2.18) we obtain

$$|\langle \alpha_\lambda^+, \boldsymbol{\psi}\rangle| \lesssim \big\|\boldsymbol{\psi} - \frac{\lambda' - b}{\lambda}(\Lambda W_\lambda, 0)\big\|_{\mathcal{E}} \ll \mathrm{e}^{-\frac{1}{2}\kappa|t|}. \tag{3.22}$$

From (2.15) we obtain $\|f(\varphi+g) - f(\varphi) - f'(\varphi)g\|_{L^{\frac{3}{2}}} \lesssim \|g\|_{\dot H^1}^2$. From (2.14) and (2.11) we have $\|(f'(\varphi) - f'(W_\mu) - f'(W_\lambda))g\|_{L^{\frac{3}{2}}} \lesssim \|\varphi - W_\mu - W_\lambda\|_{L^3}\cdot\|g\|_{L^3} \lesssim \mathrm{e}^{-\frac{3}{2}\kappa|t|}\|g\|_{\dot H^1}$. Taking the sum we obtain

$$\|f(\varphi+g) - f(\varphi) - (f'(W_\mu) + f'(W_\lambda))g\|_{L^{\frac{3}{2}}} \lesssim \|g\|_{\dot H^1}^2 + \mathrm{e}^{-\frac{3}{2}\kappa|t|}\|g\|_{\dot H^1}. \tag{3.23}$$

But $|\langle \mathcal{Y}_{\underline{\lambda}}, f'(W_\mu)g\rangle| \lesssim \|\mathcal{Y}_{\underline{\lambda}}\|_{L^{\frac{3}{2}}}\cdot\|f'(W_\mu)\|_{L^\infty}\cdot\|g\|_{L^3} \lesssim \lambda\|g\|_{\dot H^1}$, hence

$$|\langle \mathcal{Y}_{\underline{\lambda}}, f(\varphi+g) - f(\varphi) - f'(W_\lambda)g\rangle| \lesssim \frac{1}{\lambda}(\|g\|_{\dot H^1}^2 + \mathrm{e}^{-\frac{3}{2}\kappa|t|}\|g\|_{\dot H^1}) \ll \mathrm{e}^{-\frac{1}{2}\kappa|t|}. \tag{3.24}$$

Combining (3.20) with (3.21), (3.22) and (3.24) we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}a_1^+(t) = \langle \alpha_\lambda^+, J\circ \mathrm{D}^2E(\boldsymbol{W}_\lambda)\boldsymbol{g}\rangle + o(\mathrm{e}^{-\frac{1}{2}\kappa|t|}) = \frac{\nu}{\lambda}a_1^+(t) + o(\mathrm{e}^{-\frac{1}{2}\kappa|t|}),$$

where in the last step we use (3.4). This proves (3.17).

Similarly, we have

$$\big|\frac{\mathrm{d}}{\mathrm{d}t}a_1^-(t) + \frac{\nu}{\lambda(t)}a_1^-(t)\big| \le c\cdot\mathrm{e}^{-\frac{1}{2}\kappa|t|}. \tag{3.25}$$

Inequality (3.16) implies that (3.18) holds for t in a neighborhood of T. Suppose that $T_2 \in (T, T_1)$ is the last time such that (3.18) holds for $t\in[T, T_2]$. But (3.25) implies that $\frac{\mathrm{d}}{\mathrm{d}t}a_1^-(T_2)$ and $a_1^-(T_2)$ have opposite signs. Hence (3.18) cannot break down at $t = T_2$. The contradiction shows that (3.18) holds for $t\in[T, T_1]$.

In order to prove (3.19), it is more convenient to work with $\boldsymbol{h}(t)$, which was defined right before (3.8), than with $\boldsymbol{g}(t)$. We will prove the bound for $a_2^+(t)$. The proof for $a_2^-(t)$ is exactly the same. Let $\widetilde a(t) := \langle \alpha_{\mu(t)}^+, \boldsymbol{h}(t)\rangle$. We have

$$\begin{aligned}|\langle \mathcal{Y}_{\underline{\mu}}, \Lambda W_{\underline{\lambda}}\rangle| &\lesssim \|\Lambda W_{\underline{\lambda}}\|_{L^1(|x|\le 1)} + \|\Lambda W_{\underline{\lambda}}\|_{L^\infty(|x|\ge 1)} \\ &\lesssim \lambda^3\int_0^{\frac{1}{\lambda}} r^{-4}r^5\,\mathrm{d}r + \frac{1}{\lambda^3}\cdot\big(\frac{1}{\lambda}\big)^{-4} \lesssim \lambda.\end{aligned} \tag{3.26}$$

Together with (2.11) this yields

$$|a_2^+(t) - \widetilde a(t)| = |\langle \alpha_\mu^+, \boldsymbol{g}(t) - \boldsymbol{h}(t)\rangle| \lesssim b|\langle \mathcal{Y}_{\underline{\mu}}, \Lambda W_{\underline{\lambda}}\rangle| + \|S\|_{\dot H^1} \le \frac{1}{2}\mathrm{e}^{-\frac{3}{2}\kappa|t|},$$

hence it suffices to show that

$$|\widetilde{a}(t)| \leq \frac{1}{2}\mathrm{e}^{-\frac{3}{2}\kappa|t|}. \tag{3.27}$$

As in the case of $a_1^+(t)$, using (3.8) we obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{a}(t) = \big\langle -\frac{\mu'}{\mu}(\Lambda_{\mathcal{E}^*}\alpha^+)_\mu, \boldsymbol{h}\big\rangle + \langle \alpha_\mu^+, J \circ \mathrm{D}^2 E(\boldsymbol{W}_\mu)h\rangle \\ + \langle \alpha_\mu^+, (\mu'\Lambda W_{\underline{\mu}} + \lambda'\Lambda W_{\underline{\lambda}}, f(W_\mu + W_\lambda + h) - f(W_\mu) - f(W_\lambda) - f'(W_\mu)h)\rangle\end{aligned} \tag{3.28}$$

But

$$\begin{aligned}\langle \mathcal{Y}_{\underline{\mu}}, \Lambda W_{\underline{\mu}}\rangle &= 0, \\ |\langle \mathcal{Y}_{\underline{\mu}}, \Lambda W_{\underline{\lambda}}\rangle| &\lesssim \mathrm{e}^{-\kappa|t|}, \qquad \text{see (3.26)}, \\ |\langle \mathcal{Y}_{\underline{\mu}}, f(W_\mu + W_\lambda + h) - f(W_\mu + W_\lambda) - f'(W_\mu + W_\lambda)h\rangle| &\lesssim \|h\|_{\dot H^1}^2 \lesssim \mathrm{e}^{-2\kappa|t|}, \\ |\langle \mathcal{Y}_{\underline{\mu}}, f(W_\mu + W_\lambda) - f(W_\mu) - f(W_\lambda)\rangle| &= 2|\langle \mathcal{Y}_{\underline{\mu}}, W_\mu \cdot W_\lambda\rangle| \\ &= 2|\langle \mathcal{Y}_{\underline{\mu}} \cdot W_\mu, W_\lambda\rangle| \lesssim \mathrm{e}^{-2\kappa|t|}, \qquad \text{see (3.26)}, \\ |\langle \mathcal{Y}_{\underline{\mu}}, (f'(W_\mu + W_\lambda) - f'(W_\mu))h\rangle| &\lesssim \|\mathcal{Y}_{\underline{\mu}}\|_{L^6} \cdot \|W_\lambda\|_{L^2} \cdot \|h\|_{L^3} \lesssim \lambda\|h\|_{\dot H^1} \lesssim \mathrm{e}^{-2\kappa|t|},\end{aligned}$$

hence (3.28) yields

$$\big|\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{a}(t) - \frac{\nu}{\mu(t)}\widetilde{a}(t)\big| \lesssim \mathrm{e}^{-2\kappa|t|} \leq c \cdot \mathrm{e}^{-\frac{3}{2}\kappa|t|},$$

with a constant c arbitrarily small. Using (2.7), we get

$$\big|\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{a}(t)\big| \leq \frac{9\nu}{8}|\widetilde{a}(t)| + c \cdot \mathrm{e}^{-\frac{3}{2}\kappa|t|}. \tag{3.29}$$

As in the proof of (3.25), suppose that $T_2 \in (T, T_1)$ is the last time such that (3.27) holds for $t \in [T, T_2]$. This implies that $|\widetilde{a}(T_2)| = \frac{1}{2}\mathrm{e}^{-\frac{3}{2}\kappa|T_2|}$ and $|\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{a}(T_2)| \geq \frac{3}{4}\kappa \cdot \mathrm{e}^{-\frac{3}{2}\kappa|T_2|}$, thus (3.29) yields

$$\frac{3}{4}\kappa \cdot \mathrm{e}^{-\frac{3}{2}\kappa|T_2|} \leq \frac{9\nu}{16}\mathrm{e}^{-\frac{3}{2}\kappa|T_2|} + c \cdot \mathrm{e}^{-\frac{3}{2}\kappa|T_2|}.$$

But (2.1) and (3.1) give $\frac{3}{4}\kappa > \frac{9\nu}{16}$. Since c is arbitrarily small, we obtain a contradiction. □

3.3 Coercivity

Recall that we denote $L := -\Delta - f'(W)$.

Lemma 3.8. *There exist constants $c, C > 0$ such that*

- *for all $g \in \dot H^1$ radially symmetric there holds*

$$\langle g, Lg\rangle = \int_{\mathbb{R}^6} |\nabla g|^2\,\mathrm{d}x - \int_{\mathbb{R}^N} f'(W)|g|^2\,\mathrm{d}x \geq c\int_{\mathbb{R}^6} |\nabla g|^2\,\mathrm{d}x - C\big(\langle \mathcal{Z}, g\rangle^2 + \langle \mathcal{Y}, g\rangle^2\big),$$

- *if $r_1 > 0$ is large enough, then for all $g \in \dot H^1_{\mathrm{rad}}$ there holds*

$$(1-2c)\int_{|x|\leq r_1} |\nabla g|^2\,\mathrm{d}x + c\int_{|x|\geq r_1} |\nabla g|^2\,\mathrm{d}x - \int_{\mathbb{R}^6} f'(W)|g|^2\,\mathrm{d}x \geq -C\big(\langle \mathcal{Z}, g\rangle^2 + \langle \mathcal{Y}, g\rangle^2\big), \tag{3.30}$$

- *if* $r_2 > 0$ *is small enough, then for all* $g \in \dot H^1_{\mathrm{rad}}$ *there holds*

$$(1-2c)\int_{|x|\geq r_2}|\nabla g|^2\,\mathrm{d}x + c\int_{|x|\leq r_2}|\nabla g|^2\,\mathrm{d}x - \int_{\mathbb{R}^6} f'(W)|g|^2\,\mathrm{d}x \geq -C\big(\langle \mathcal{Z}, g\rangle^2 + \langle \mathcal{Y}, g\rangle^2\big). \tag{3.31}$$

Proof. This is exactly Lemma 2.1 in [41], see also [58, Lemma 2.1]. □

Lemma 3.9. *There exists a constant* $\eta > 0$ *such that if* $\frac{\lambda}{\mu} < \eta$ *and* $\|\boldsymbol{U} - (\boldsymbol{W}_\mu + \boldsymbol{W}_\lambda)\|_{\mathcal{E}} < \eta$, *then for all* $\boldsymbol{g} \in \mathcal{E}$ *there holds*

$$\frac12\langle \mathrm{D}^2E(\boldsymbol{U})\boldsymbol{g},\boldsymbol{g}\rangle + 2\big(\langle\alpha_\lambda^-,\boldsymbol{g}\rangle^2 + \langle\alpha_\lambda^+,\boldsymbol{g}\rangle^2 + \langle\frac{1}{\lambda}\mathcal{Z}_{\underline{\lambda}}, g\rangle^2 + \langle\alpha_\mu^-,\boldsymbol{g}\rangle^2 + \langle\alpha_\mu^+,\boldsymbol{g}\rangle^2 + \langle\frac{1}{\mu}\mathcal{Z}_{\underline{\mu}}, g\rangle^2\big) \gtrsim \|\boldsymbol{g}\|_{\mathcal{E}}^2.$$

Proof. We will repeat with minor changes the proof of [41, Lemma 3.5].

Step 1 Without loss of generality we can assume that $\mu = 1$. Consider the operator $\boldsymbol{L}_\lambda$ defined by the following formula:

$$\boldsymbol{L}_\lambda := \begin{pmatrix} -\Delta - f'(W_\lambda) - f'(W) & 0 \\ 0 & \mathrm{Id}\end{pmatrix}.$$

From the fact that $\|f'(U) - f'(W) - f'(W_\lambda)\|_{L^3} \lesssim \|U - (W + W_\lambda)\|_{L^3}$ we obtain

$$|\langle \mathrm{D}^2E(\boldsymbol{U})\boldsymbol{g},\boldsymbol{g}\rangle - \langle \boldsymbol{L}_\lambda\boldsymbol{g},\boldsymbol{g}\rangle| \leq c\|\boldsymbol{g}\|_{\mathcal{E}}^2, \qquad \forall \boldsymbol{g}\in\mathcal{E}, \tag{3.32}$$

with $c > 0$ small when η and λ_0 are small.

Step 2. In view of (3.32), it suffices to prove that if $\lambda < \lambda_0$, then

$$\frac12\langle \boldsymbol{L}_\lambda\boldsymbol{g},\boldsymbol{g}\rangle + 2\big(\langle\alpha_{\lambda_1}^-,\boldsymbol{g}\rangle^2 + \langle\alpha_{\lambda_1}^+,\boldsymbol{g}\rangle^2 + \langle\alpha_{\lambda_2}^-,\boldsymbol{g}\rangle^2 + \langle\alpha_{\lambda_2}^+,\boldsymbol{g}\rangle^2 + \langle\frac{1}{\lambda}\mathcal{Z}_{\underline{\lambda}}, g\rangle^2 + \langle \mathcal{Z}, g\rangle^2\big) \gtrsim \|\boldsymbol{g}\|_{\mathcal{E}}^2.$$

Let $a_1^- := \langle\alpha_\lambda^-,\boldsymbol{g}\rangle$, $a_1^+ := \langle\alpha_\lambda^+,\boldsymbol{g}\rangle$, $a_2^- := \langle\alpha^-,\boldsymbol{g}\rangle$, $a_2^+ := \langle\alpha^+,\boldsymbol{g}\rangle$, $b_1 := \langle\mathcal{Z},\Lambda W\rangle^{-1}\cdot\langle\frac{1}{\lambda}\mathcal{Z}_{\underline{\lambda}}, g\rangle$, $b_2 := \langle\mathcal{Z},\Lambda W\rangle^{-1}\cdot\langle\mathcal{Z}, g\rangle$ and decompose

$$\boldsymbol{g} = a_1^-\mathcal{Y}_\lambda^- + a_1^+\mathcal{Y}_\lambda^+ + a_2^-\mathcal{Y}^- + a_2^+\mathcal{Y}^+ + b_1\Lambda\boldsymbol{W}_\lambda + b_2\Lambda\boldsymbol{W} + \boldsymbol{k}.$$

Using the fact that

$$|\langle\alpha^\pm,\mathcal{Y}_\lambda^\pm\rangle| + |\langle\alpha_\lambda^\pm,\mathcal{Y}^\pm\rangle| + |\langle\frac{1}{\lambda}\mathcal{Z}_{\underline{\lambda}},\mathcal{Y}\rangle| + |\langle\mathcal{Z},\mathcal{Y}_\lambda\rangle| \lesssim \lambda^2,$$
$$|a_1^-| + |a_1^+| + |a_2^-| + |a_2^+| + |b_1| + |b_2| \lesssim \|\boldsymbol{g}\|_{\mathcal{E}},$$
$$\langle\alpha^-,\mathcal{Y}^+\rangle = \langle\alpha^+,\mathcal{Y}^-\rangle = \langle\mathcal{Z},\mathcal{Y}\rangle = \langle\mathcal{Y},\Lambda W\rangle = 0$$

we obtain

$$\langle\alpha^-,\boldsymbol{k}\rangle^2 + \langle\alpha^+,\boldsymbol{k}\rangle^2 + \langle\alpha_\lambda^-,\boldsymbol{k}\rangle^2 + \langle\alpha_\lambda^+,\boldsymbol{k}\rangle^2 + \langle\mathcal{Z},k\rangle^2 + \langle\frac{1}{\lambda}\mathcal{Z}_{\underline{\lambda}},k\rangle^2 \lesssim \lambda^4\cdot\|\boldsymbol{g}\|_{\mathcal{E}}^2. \tag{3.33}$$

Since $\boldsymbol{L}_\lambda$ is self-adjoint, we can write

$$\begin{aligned}\frac12\langle\boldsymbol{L}_\lambda\boldsymbol{g},\boldsymbol{g}\rangle &= \frac12\langle\boldsymbol{L}_\lambda\boldsymbol{k},\boldsymbol{k}\rangle \\ &\quad + \langle\boldsymbol{L}_\lambda(a_2^-\mathcal{Y}^- + a_2^+\mathcal{Y}^+ + b_2\Lambda\boldsymbol{W}),\boldsymbol{k}\rangle + \langle\boldsymbol{L}_\lambda(a_1^-\mathcal{Y}_\lambda^- + a_1^+\mathcal{Y}_\lambda^+ + b_1\Lambda\boldsymbol{W}_\lambda),\boldsymbol{k}\rangle \\ &\quad + \frac12\langle\boldsymbol{L}_\lambda(a_2^-\mathcal{Y}^- + a_2^+\mathcal{Y}^+ + b_2\Lambda\boldsymbol{W}), a_2^-\mathcal{Y}^- + a_2^+\mathcal{Y}^+ + b_2\Lambda\boldsymbol{W}\rangle \\ &\quad + \frac12\langle\boldsymbol{L}_\lambda(a_1^-\mathcal{Y}_\lambda^- + a_1^+\mathcal{Y}_\lambda^+ + b_1\Lambda\boldsymbol{W}_\lambda), a_1^-\mathcal{Y}_\lambda^- + a_1^+\mathcal{Y}_\lambda^+ + b_1\Lambda\boldsymbol{W}_\lambda\rangle \\ &\quad + \langle\boldsymbol{L}_\lambda(a_2^-\mathcal{Y}^- + a_2^+\mathcal{Y}^+ + b_2\Lambda\boldsymbol{W}), a_1^-\mathcal{Y}_\lambda^- + a_1^+\mathcal{Y}_\lambda^+ + b_1\Lambda\boldsymbol{W}_\lambda\rangle.\end{aligned} \tag{3.34}$$

It is easy to see that $\|f'(W)\mathcal{Y}_\lambda\|_{L^{\frac{3}{2}}} \to 0$, $\|f'(W)\Lambda W_\lambda\|_{L^{\frac{3}{2}}} \to 0$, $\|f'(W_\lambda)\mathcal{Y}\|_{L^{\frac{3}{2}}} \to 0$ and $\|f'(W_\lambda)\Lambda W\|_{L^{\frac{3}{2}}} \to 0$ as $\lambda \to 0$. This and (3.2), (3.3) imply

$$\begin{gathered}\|\boldsymbol{L}_\lambda\mathcal{Y}^- + 2\alpha^+\|_{\mathcal{E}^*} + \|\boldsymbol{L}_\lambda\mathcal{Y}^+ + 2\alpha^-\|_{\mathcal{E}^*} + \|\boldsymbol{L}_\lambda\Lambda\boldsymbol{W}\|_{\mathcal{E}^*}\\ +\|\boldsymbol{L}_\lambda\mathcal{Y}_\lambda^- + 2\alpha_\lambda^+\|_{\mathcal{E}^*} + \|\boldsymbol{L}_\lambda\mathcal{Y}_\lambda^+ + 2\alpha_\lambda^-\|_{\mathcal{E}^*} + \|\boldsymbol{L}_\lambda\Lambda\boldsymbol{W}_\lambda\|_{\mathcal{E}^*} \xrightarrow[\lambda\to 0]{} 0.\end{gathered}$$

Plugging this into (3.34) and using (3.33) we obtain

$$\frac{1}{2}\langle\boldsymbol{L}_\lambda\boldsymbol{g},\boldsymbol{g}\rangle \geq -2a_2^-a_2^+ - 2a_1^-a_1^+ + \frac{1}{2}\langle\boldsymbol{L}_\lambda\boldsymbol{k},\boldsymbol{k}\rangle - \widetilde{c}\|\boldsymbol{g}\|_{\mathcal{E}}^2, \tag{3.35}$$

where $\widetilde{c} \to 0$ as $\lambda \to 0$.

Applying (3.30) with $r_1 = \lambda^{-\frac{1}{2}}$, rescaling and using (3.33) we get, for λ small enough,

$$(1-2c)\int_{|x|\leq\sqrt{\lambda}}|\nabla k|^2\,\mathrm{d}x + c\int_{|x|\geq\sqrt{\lambda}}|\nabla k|^2\,\mathrm{d}x - \int_{\mathbb{R}^6} f'(W_\lambda)|k|^2\,\mathrm{d}x \geq -\widetilde{c}\|\boldsymbol{g}\|_{\mathcal{E}}^2. \tag{3.36}$$

From (3.31) with $r_2 = \sqrt{\lambda}$ we have

$$(1-2c)\int_{|x|\geq\sqrt{\lambda}}|\nabla k|^2\,\mathrm{d}x + c\int_{|x|\leq\sqrt{\lambda}}|\nabla k|^2\,\mathrm{d}x - \int_{\mathbb{R}^6} f'(W)|k|^2\,\mathrm{d}x \geq -\widetilde{c}\|\boldsymbol{g}\|_{\mathcal{E}}^2. \tag{3.37}$$

Taking the sum of (3.36) and (3.37), and using (3.35), we obtain

$$\frac{1}{2}\langle\boldsymbol{L}_\lambda\boldsymbol{g},\boldsymbol{g}\rangle \geq -2a_2^-a_2^+ - 2a_1^-a_1^+ + c\|\boldsymbol{k}\|_{\mathcal{E}}^2 - 2\widetilde{c}\|\boldsymbol{g}\|_{\mathcal{E}}^2.$$

The conclusion follows if we take $\widetilde{c}$ small enough. □

3.4 Definition of the mixed energy-virial functional

Lemma 3.10. *For any $c > 0$ and $R > 0$ there exists a radial function $q(x) = q_{c,R}(x) \in C^{3,1}(\mathbb{R}^6)$ with the following properties:*

(P1) $q(x) = \frac{1}{2}|x|^2$ for $|x| \leq R$,

(P2) there exists $\widetilde{R} > 0$ (depending on c and R) such that $q(x) \equiv \mathrm{const}$ for $|x| \geq \widetilde{R}$,

(P3) $|\nabla q(x)| \lesssim |x|$ and $|\Delta q(x)| \lesssim 1$ for all $x \in \mathbb{R}^6$, with constants independent of c and R,

(P4) $\sum_{1\leq i,j\leq 6}\big(\partial_{x_ix_j}q(x)\big)v_iv_j \geq -c\sum_{i=1}^6 v_i^2$, for all $x \in \mathbb{R}^6, v_i \in \mathbb{R}$,

(P5) $\Delta^2 q(x) \leq c\cdot|x|^{-2}$, for all $x \in \mathbb{R}^6$.

Remark 3.11. We require $C^{3,1}$ regularity in order not to worry about boundary terms in Pohozaev identities, see the proof of (3.41).

Proof. It suffices to prove the result for $R = 1$ since the function $q_R(x) := R^2q(\frac{x}{R})$ satisfies the listed properties if and only if $q(x)$ does.

Let r denote the radial coordinate. Define $q_0(x)$ by the formula

$$q_0(r) := \begin{cases}\frac{1}{2}\cdot r^2 & r\leq 1\\ \frac{8}{5}\cdot r - \frac{3}{2} + \frac{1}{2}\cdot r^{-2} - \frac{1}{10}\cdot r^{-4} & r\geq 1.\end{cases}$$

A direct computation shows that for $r > 1$ we have $q_0'(r) = \frac{8}{5} - r^{-3} + \frac{2}{5}r^{-5}$, $q_0''(r) = 3r^{-4} - 2r^{-6} > 0$ (so $q_0(x)$ is convex), $q_0'''(r) = 12(-r^{-5} + r^{-7})$ and $\Delta^2 q_0(r) = -24r^{-3}$. Hence q_0 satisfies all the listed properties except for (P2). We correct it as follows.

Let $e_j(r) := \frac{1}{j!}r^j \cdot \chi(r)$ for $j \in \{1,2,3\}$ and let $R_0 \gg 1$. We define

$$q(r) := \begin{cases} q_0(r) & r \leq R_0 \\ q_0(R_0) + \sum_{j=1}^3 q_0^{(j)}(R_0)\cdot R_0^j \cdot e_j(-1 + R_0^{-1}r) & r \geq R_0. \end{cases}$$

Note that $q_0'(R_0) \sim 1$, $q_0''(R_0) \sim R_0^{-4}$ and $q_0'''(R_0) \sim R_0^{-5}$. It is clear that $q(x) \in C^{3,1}(\mathbb{R}^6)$. Property (P1) holds since $R_0 > 1$. By the definition of the functions e_j we have $q(r) = q_0(R_0) = \text{const}$ for $r \geq 3R_0$, hence (P2) holds with $\widetilde{R} = 3R_0$. From the definition of $q(r)$ we get $|\partial_r q(r)| \lesssim 1$ and $|\partial_r^2 q(r)| \lesssim R_0^{-4}$ for $r \geq R_0$, with a constant independent of R_0, which implies (P3). Similarly, $|\partial_{x_i x_j} q(x)| \lesssim R_0^{-1}$ for $|x| \geq R_0$, which implies (P4) if R_0 is large enough. Finally $|\Delta^2 q(x)| \lesssim R_0^{-3}$ for $|x| \geq R_0$ and $\Delta^2 q(x) = 0$ for $|x| \geq 3R_0$. This proves (P5) if R_0 is large enough. □

In the sequel $q(x)$ always denotes a function of class $C^{3,1}(\mathbb{R}^6)$ verifying (P1)–(P5) with sufficiently small c and sufficiently large R.

For $\lambda > 0$ we define the operators $A(\lambda)$ and $A_0(\lambda)$ as follows:

$$\begin{aligned} [A(\lambda)h](x) &:= \frac{1}{3\lambda}\Delta q(\frac{x}{\lambda})h(x) + \nabla q(\frac{x}{\lambda})\cdot\nabla h(x), \\ [A_0(\lambda)h](x) &:= \frac{1}{2\lambda}\Delta q(\frac{x}{\lambda})h(x) + \nabla q(\frac{x}{\lambda})\cdot\nabla h(x). \end{aligned} \tag{3.38}$$

Combining these definitions with the fact that $q(x)$ is an approximation of $\frac{1}{2}|x|^2$ we see that $A(\lambda)$ and $A_0(\lambda)$ are approximations (in a sense not yet precised) of $\frac{1}{\lambda}\Lambda$ and $\frac{1}{\lambda}\Lambda_0$ respectively. We will write A and A_0 instead of $A(1)$ and $A_0(1)$ respectively. Note the following scale-change formulas, which follow directly from the definitions:

$$\forall h \in \dot{H}^1: \qquad A(\lambda)(h_\lambda) = (Ah)_{\underline{\lambda}}, \quad A_0(\lambda)(h_\lambda) = (A_0 h)_{\underline{\lambda}}. \tag{3.39}$$

Lemma 3.12. *The operators $A(\lambda)$ and $A_0(\lambda)$ have the following properties:*

- *the families $\{A(\lambda) : \lambda > 0\}$, $\{A_0(\lambda) : \lambda > 0\}$, $\{\lambda\partial_\lambda A(\lambda) : \lambda > 0\}$ and $\{\lambda\partial_\lambda A_0(\lambda) : \lambda > 0\}$ are bounded in $\mathscr{L}(\dot{H}^1; L^2)$, with the bound depending on the choice of the function $q(x)$,*

- *for all $h_1, h_2 \in X^1$ and $\lambda > 0$ there holds*

$$\langle A(\lambda)h_1, f(h_1 + h_2) - f(h_1) - f'(h_1)h_2\rangle = -\langle A(\lambda)h_2, f(h_1 + h_2) - f(h_1)\rangle, \tag{3.40}$$

- *for any $c_0 > 0$, if we choose c in Lemma 3.10 small enough, then for all $h \in X^1 \times H^1$ there holds*

$$\langle A_0(\lambda)h, \Delta h\rangle \leq \frac{c_0}{\lambda}\|h\|_{\dot{H}^1}^2 - \frac{1}{\lambda}\int_{|x|\leq R\lambda}|\nabla h(x)|^2\,\mathrm{d}x. \tag{3.41}$$

- *assuming (2.7) and (3.10), for any $c_0 > 0$ there holds*

$$\|\Lambda_0\Lambda W_{\underline{\lambda(t)}} - A_0(\lambda(t))\Lambda W_{\lambda(t)}\|_{L^2} \leq c_0, \tag{3.42}$$

$$\|\dot{\varphi}(t) + b(t)\cdot A(\lambda(t))\varphi(t)\|_{L^3} \leq c_0, \tag{3.43}$$

$$\Big|\int\frac{1}{6}\Delta q(\frac{x}{\lambda})(f(\varphi + g) - f(\varphi))g\,\mathrm{d}x - \int f'(W_\lambda)g^2\,\mathrm{d}x\Big| \leq c_0 C_0^2 \mathrm{e}^{-3\kappa|t|}. \tag{3.44}$$

provided that the constant R in the definition of $q(x)$ is chosen large enough.

Proof. Since $\nabla q(x)$ and $\nabla^2 q(x)$ are continuous and of compact support, it is clear that A and A_0 are bounded operators $\dot H^1 \to L^2$. From the invariance (3.39) we see that $A(\lambda)$ and $A_0(\lambda)$ have the same norms as A and A_0 respectively. For $\lambda\partial_\lambda A(\lambda)$ and $\lambda\partial_\lambda A_0(\lambda)$ the proof is similar. We compute

$$\partial_\lambda A(\lambda) = -\frac{1}{3\lambda^2}\Delta q\big(\frac{x}{\lambda}\big) - \frac{1}{3\lambda^3}x\cdot\nabla\Delta q\big(\frac{x}{\lambda}\big) - \frac{1}{\lambda^2}x\cdot\nabla^2 q\big(\frac{x}{\lambda}\big)\cdot\nabla.$$

Since $\nabla q(x)$, $\nabla^2 q(x)$ and $\nabla^3 q(x)$ are continuous and of compact support, boundedness follows.

In (3.40) both sides are continuous for the X^1 topology, hence we may assume that $h_1, h_2 \in C_0^\infty$. We may also assume without loss of generality that $\lambda = 1$. Observe that for any $h \in C_0^\infty$ there holds $h\cdot f(h) = 3\cdot F(h)$ and $\nabla h\cdot f(h) = \nabla F(h)$, hence

$$\langle Ah, f(h)\rangle = \int \big(\frac{1}{3}\Delta q\cdot h + \nabla q\cdot\nabla h\big)f(h)\,\mathrm{d}x = \int \Delta q\cdot F(h) + \nabla q\cdot\nabla F(h)\,\mathrm{d}x = 0.$$

Using this for $h = h_1 + h_2$ and for $h = h_1$, (3.40) is seen to be equivalent to

$$\langle Ah_2, f(h_1)\rangle + \langle Ah_1, f'(h_1)h_2\rangle = 0. \tag{3.45}$$

Expanding the left side using the definition of A we obtain

$$\begin{aligned}\langle Ah_2, f(h_1)\rangle + \langle Ah_1, f'(h_1)h_2\rangle &= \int \frac{1}{3}\Delta q\cdot h_2\cdot f(h_1) + \nabla q\cdot\nabla h_2\cdot f(h_1)\,\mathrm{d}x \\ &+ \int \frac{1}{3}\Delta q\cdot h_1\cdot f'(h_1)\cdot h_2 + \nabla q\cdot\nabla h_1\cdot f'(h_1)\cdot h_2\,\mathrm{d}x\end{aligned}$$

Integrating by parts the term containing ∇h_2 and using the formulas $h_1\cdot f'(h_1) = 2f(h_1)$ and $\nabla h_1\cdot f'(h_1) = \nabla f(h_1)$, this can be rewritten as

$$\big\langle h_2, \frac{1}{3}\Delta q\cdot f(h_1) - \Delta q\cdot f(h_1) - \nabla q\cdot\nabla f(h_1) + \frac{2}{3}\Delta q\cdot f(h_1) + \nabla q\cdot\nabla f(h_1)\big\rangle = 0,$$

which proves (3.45).

Inequality (3.41) follows easily from (P1), (P4) and (P5), once we check the following identity (valid in any dimension N, and used here for $N = 6$):

$$\begin{aligned}&\int \Delta h(x)\cdot\big(\frac{1}{2\lambda}\Delta q\big(\frac{x}{\lambda}\big)h(x) + \nabla q\big(\frac{x}{\lambda}\big)\cdot\nabla h(x)\big)\,\mathrm{d}x \\ &= -\frac{1}{4\lambda}\int (\Delta^2 q)\big(\frac{x}{\lambda}\big)h(x)^2\,\mathrm{d}x - \frac{1}{\lambda}\int\sum_{i,j=1}^N \partial_{ij}q\big(\frac{x}{\lambda}\big)\partial_i h(x)\partial_j h(x)\,\mathrm{d}x.\end{aligned} \tag{3.46}$$

Without loss of generality we can assume that $\lambda = 1$ (it suffices to replace q by its rescaled version). By a density argument, we can also assume that $q, h \in C_0^\infty$ (we use here the fact that $q \in C^{3,1}$), and (3.46) follows from integration by parts:

$$\begin{aligned}
&\int \frac{1}{2}\Delta h\cdot \Delta q\cdot h + \Delta h\cdot \nabla q\cdot \nabla h\,\mathrm{d}x = \int \sum_{i,j=1}^N \big(\frac{1}{2}\partial_{ii}h\cdot\partial_{jj}q\cdot h + \partial_{ii}h\cdot \partial_j q\cdot \partial_j h\big)\,\mathrm{d}x \\
&= \int -\frac{1}{2}\sum_{i,j}\partial_i h(\partial_{jj}q\partial_i h + \partial_{ijj}q\cdot h) + \sum_i \frac{1}{2}\partial_i((\partial_i h)^2)\partial_i q \\
&+ \sum_{i\neq j}\big(-\frac{1}{2}\partial_j(\partial_i h)^2\partial_j q - \partial_{ij}q\partial_i h\partial_j h\big)\,\mathrm{d}x \\
&= \int -\frac{1}{2}\sum_{i,j}\big(\partial_{jj}q(\partial_i h)^2 + \frac{1}{2}\partial_{iijj}q\cdot h^2\big) - \frac{1}{2}\sum_i \partial_{ii}q(\partial_i h)^2 \\
&+ \frac{1}{2}\sum_{i\neq j}\partial_{jj}q(\partial_i h)^2 - \sum_{i\neq j}\partial_{ij}q\partial_i h\partial_j h\,\mathrm{d}x \\
&= \int -\frac{1}{4}\sum_{i,j}\partial_{iijj}q\cdot h^2 - \sum_{i,j}\partial_{ij}q\partial_i h\partial_j h\,\mathrm{d}x.
\end{aligned}$$

Estimate (3.42) is invariant by rescaling, hence we can assume that $\lambda = 1$. For $|x| \le R$ we have $A_0\Lambda W(x) = \Lambda_0\Lambda W(x)$. From (P3) in Lemma 3.10 we get $|A_0\Lambda W(x)| + |\Lambda_0\Lambda W(x)| \lesssim |x|^{-4}$ for $|x| \ge R$, with a constant independent of R. Thus $\|\Lambda_0\Lambda W - A_0\Lambda W\|_{L^2} \le c_0$ if R is large enough.

A similar reasoning yields $\|\Lambda W_{\underline{\lambda}} - A(\lambda)W_\lambda\|_{L^3} \ll \lambda^{-1}$ as $R \to +\infty$. Since $b(t) \sim \lambda(t)$, this gives

$$\|\dot\varphi + bA(\lambda)W_\lambda\|_{L^3} \le \frac{c_0}{3}, \qquad \text{if } R \text{ is large enough.} \tag{3.47}$$

From (P2) in Lemma 3.10 it follows that $\operatorname{supp}(A(\lambda)W_\mu) \subset B(0, \widetilde{R}\cdot\lambda)$. Since $\|A(\lambda)W_\mu\|_{L^\infty} \lesssim_q \frac{1}{\lambda}$, we have

$$\|bA(\lambda)W_\mu\|_{L^3} \le \frac{c_0}{3}, \qquad \text{if } |T_0| \text{ is large enough.} \tag{3.48}$$

To finish the proof, we have to check that

$$\big\|bA(\lambda)\big(\chi\cdot\big(\lambda(t)^2P_{\lambda(t)} + b(t)^2Q_{\lambda(t)}\big)\big)\big\|_{L^3} \le \frac{c_0}{3}, \qquad \text{if } |T_0| \text{ is large enough.} \tag{3.49}$$

We have the bound $\|(\chi + |\nabla\chi|)P_\lambda\|_{L^3} \lesssim \int_0^{\frac{2}{\lambda}}\frac{1}{r^6}r^5\,\mathrm{d}r \lesssim |\log\lambda|$ (similarly with Q_λ), hence $\|(\chi+|\nabla\chi|)(\lambda^2P_\lambda + b^2Q_\lambda)\|_{L^3} \lesssim (\lambda^2+b^2)|\log\lambda| \ll 1$ as $|T_0| \to +\infty$. We have also $\|\nabla(\lambda^2P_\lambda + b^2Q_\lambda)\|_{L^3} \lesssim \frac{1}{\lambda}(\lambda^2+b^2) \ll 1$. Since q is smooth and constant at infinity, we have $\big|b\cdot\big(\frac{1}{\lambda}\Delta q\big(\frac{x}{\lambda}\big) + \nabla q\big(\frac{x}{\lambda}\big)\big)\big| \lesssim 1$. The constant depends on the choice of the function q, but this is not a concern here. We obtain

$$\begin{aligned}
&\big\|bA(\lambda)\big(\chi\cdot\big(\lambda(t)^2P_{\lambda(t)} + b(t)^2Q_{\lambda(t)}\big)\big)\big\|_{L^3} \lesssim \big|b\cdot\big(\frac{1}{\lambda}\Delta q\big(\frac{x}{\lambda}\big) + \nabla q\big(\frac{x}{\lambda}\big)\big)\big|\cdot \\
&\cdot\big(\|(\chi + |\nabla\chi|)(\lambda^2P_\lambda + b^2Q_\lambda)\|_{L^3} + \|\nabla(\lambda^2P_\lambda + b^2Q_\lambda)\|_{L^3}\big) \ll 1,
\end{aligned}$$

hence (3.49)

Putting together (3.47), (3.48) and (3.49) we get (3.43).

In order to prove (3.44), note first that boundedness of Δq and (3.23) yield

$$\Big|\int \frac{1}{6}\Delta q\big(\frac{x}{\lambda}\big)(f(\varphi+g) - f(\varphi))g\,\mathrm{d}x - \int\frac{1}{6}\Delta q\big(\frac{x}{\lambda}\big)(f'(W_\mu) + f'(W_\lambda))g^2\,\mathrm{d}x\Big| \ll \mathrm{e}^{-3\kappa|t|}.$$

Since ∇q is of compact support, we have $\left|\int \frac{1}{6}\Delta q\big(\frac{x}{\lambda}\big)W_\mu g^2\,\mathrm{d}x\right| \ll \|g\|_{\dot H^1}^2$. As in the proof of (3.43), on can show that $\big\|\frac{1}{6}\Delta q\big(\frac{x}{\lambda}\big)f'(W_\lambda) - f'(W_\lambda)\big\|_{L^3} \to 0$ as $R \to +\infty$. This finishes the proof. □

For $t \in [T, T_0]$ we define:

- the *nonlinear energy functional*

$$\begin{aligned} I(t) &:= \int \frac{1}{2}|\dot g(t)|^2 + \frac{1}{2}|\nabla g(t)|^2 - \big(F(\varphi(t)+g(t)) - F(\varphi(t)) - f(\varphi(t))g(t)\big)\,\mathrm{d}x \\ &= E(\boldsymbol{\varphi}(t)+\boldsymbol{g}(t)) - E(\boldsymbol{\varphi}(t)) - \langle \mathrm{D}E(\boldsymbol{\varphi}(t), \boldsymbol{g}(t))\rangle, \end{aligned}$$

- the *localized virial functional*

$$J(t) := \int \dot g(t)\cdot A_0(\lambda(t))g(t)\,\mathrm{d}x,$$

- the *mixed energy-virial functional*

$$H(t) := I(t) + b(t)J(t).$$

From (2.16) we have

$$\big|I(t) - \frac{1}{2}\langle \mathrm{D}^2E(\boldsymbol{\varphi}(t))\boldsymbol{g}(t), \boldsymbol{g}(t)\rangle\big| \lesssim \|\boldsymbol{g}(t)\|_{\mathcal{E}}^3.$$

Note that $H(t)$ depends on the function $q(x)$ used in the definition of $A_0(\lambda)$, see (3.38). From the first statement in Lemma 3.12 we deduce that

$$|J(t)| \lesssim_q \|\boldsymbol{g}(t)\|_{\mathcal{E}}^2,$$

where the constant in the inequality depends on the choice of the function $q(x)$. Thus (2.9) and (3.10) imply that for $t \le T_0$ with $|T_0|$ large enough there holds

$$\big|H(t) - \frac{1}{2}\langle \mathrm{D}^2E(\boldsymbol{\varphi}(t))\boldsymbol{g}(t), \boldsymbol{g}(t)\rangle\big| \le c\|\boldsymbol{g}(t)\|_{\mathcal{E}}^2, \tag{3.50}$$

with $c > 0$ as small as we wish.

3.5 Energy estimates via the mixed energy-virial functional

Lemma 3.13. *Let $c_1 > 0$. If C_0 is sufficiently large, then there exists a function $q(x)$ and $T_0 < 0$ with the following property. If $T_1 < T_0$ and* (2.7), (3.10), (3.11) *hold for $t \in [T, T_1]$, then for $t \in [T, T_1]$ there holds*

$$H'(t) \le c_1 \cdot C_0^2 \cdot \mathrm{e}^{-3\kappa|t|}. \tag{3.51}$$

This lemma is the key step in proving Proposition 3.3. We will postpone its slightly technical proof.

Proof of Proposition 3.3 assuming Lemma 3.13. We first show (3.13). From (3.15) and (3.10) we obtain

$$|\mu(t) - 1| = |\mu(t) - \mu(T)| \lesssim \int_{-\infty}^t C_0 \cdot \mathrm{e}^{-\frac{3}{2}\kappa_0|\tau|}\,\mathrm{d}\tau \lesssim C_0 \cdot \mathrm{e}^{-\frac{3}{2}\kappa_0|t|}. \tag{3.52}$$

Again from (3.15) and (3.10) we have $|\lambda'(t) - b(t)| \lesssim C_0 \cdot \mathrm{e}^{-\frac{3}{2}\kappa|t|}$. Multiplying by $b'(t) = \kappa^2 \cdot \frac{\lambda(t)}{\mu(t)^2} \sim \mathrm{e}^{-\kappa|t|}$, cf. (2.8) and (2.7), we obtain $\left|\frac{\mathrm{d}}{\mathrm{d}t}\big(b(t)^2 - \kappa^2 \cdot \frac{\lambda(t)^2}{\mu(t)^2}\big)\right| \lesssim C_0 \cdot \mathrm{e}^{-\frac{5}{2}\kappa|t|}$. Since $b(T) = \kappa \cdot \lambda(T)$ and $\mu(T) = 1$, this yields $\left|b(t)^2 - \kappa^2 \cdot \frac{\lambda(t)^2}{\mu(t)^2}\right| \lesssim C_0 \cdot \mathrm{e}^{-\frac{5}{2}\kappa|t|}$. But $b(t) + \kappa \cdot \frac{\lambda(t)}{\mu(t)} \sim \mathrm{e}^{-\kappa|t|}$, see (2.7) and (2.9), hence

$$\left|b(t) - \kappa \cdot \frac{\lambda(t)}{\mu(t)}\right| \lesssim C_0 \cdot \mathrm{e}^{-\frac{3}{2}\kappa|t|}. \tag{3.53}$$

Bound (3.52) implies that $\left|\frac{\lambda(t)}{\mu(t)} - \lambda(t)\right| \ll \mathrm{e}^{-\frac{3}{2}\kappa|t|}$, thus (3.53) yields $|\lambda'(t) - \kappa\cdot\lambda(t)| \lesssim C_0 \cdot \mathrm{e}^{-\frac{3}{2}\kappa|t|}$. Integrating and using $\lambda(T) = \frac{1}{\kappa}\mathrm{e}^{-\kappa|T|}$ we obtain (3.13).

We turn to the proof of (3.12). The initial data at $t = T$ satisfy $\|\boldsymbol{g}(T)\|_\mathcal{E} \lesssim \mathrm{e}^{-\frac{3}{2}\kappa|T|}$, thus (3.50) implies that $H(T) \lesssim \mathrm{e}^{-3\kappa|T|}$. If C_0 is large enough, then integrating (3.51) we get $H(t) \leq c \cdot C_0^2 \cdot \mathrm{e}^{-3\kappa|t|}$, with a small constant c. Now (3.50) implies

$$\langle \mathrm{D}^2 E(\boldsymbol{\varphi}(t))\boldsymbol{g}(t), \boldsymbol{g}(t)\rangle \lesssim c \cdot C_0^2 \cdot \mathrm{e}^{-3\kappa|t|}, \qquad \text{with } c \text{ small.} \tag{3.54}$$

Since $\|\boldsymbol{\varphi}(t) - \boldsymbol{W}_{\lambda(t)}\|_\mathcal{E}$ is small, Lemma 3.9 together with (3.14), (3.11), (3.18) and (3.19) yields

$$\|\boldsymbol{g}\|_\mathcal{E}^2 \lesssim (cC_0^2 + 1)\mathrm{e}^{-3\kappa|t|},$$

Eventually enlarging C_0, we obtain (3.12), if c in (3.54) is taken sufficiently small. □

Proof of Lemma 3.13. In this proof we say that a term is negligible if its contribution is $\leq c \cdot C_0^2 \cdot \mathrm{e}^{-3\kappa|t|}$. We write $\mathrm{Value}_1 \simeq \mathrm{Value}_2$ if $|\mathrm{Value}_1 - \mathrm{Value}_2| \leq c \cdot C_0^2 \cdot \mathrm{e}^{-3\kappa|t|}$. The order of choosing the parameters is the following: we will first choose $q(x)$ independently of C_0, then C_0, which may depend on $q(x)$, and finally $|T_0|$.

Step 1 (Derivative of the energy functional) Using the definition of $I(t)$, the conservation of energy, formulas (2.13), (3.9) and self-adjointness of $\mathrm{D}^2E(\boldsymbol{\varphi})$ we compute:

$$\begin{aligned}
I'(t) &= 0 - \langle \mathrm{D}E(\boldsymbol{\varphi}), \partial_t\boldsymbol{\varphi}\rangle - \langle \mathrm{D}^2E(\boldsymbol{\varphi})\partial_t\boldsymbol{\varphi}, \boldsymbol{g}\rangle - \langle \mathrm{D}E(\boldsymbol{\varphi}), \partial_t\boldsymbol{g}\rangle \\
&= -\langle \mathrm{D}E(\boldsymbol{\varphi}), J\circ \mathrm{D}E(\boldsymbol{\varphi}) + \boldsymbol{\psi}\rangle - \langle J\circ\mathrm{D}E(\boldsymbol{\varphi}) + \boldsymbol{\psi}, \mathrm{D}^2E(\boldsymbol{\varphi})\boldsymbol{g}\rangle \\
&\quad - \langle \mathrm{D}E(\boldsymbol{\varphi}), J\circ(\mathrm{D}E(\boldsymbol{\varphi}+\boldsymbol{g}) - \mathrm{D}E(\boldsymbol{\varphi})) - \boldsymbol{\psi}\rangle \\
&= -\langle \mathrm{D}^2E(\boldsymbol{\varphi})\boldsymbol{\psi}, \boldsymbol{g}\rangle - \langle \mathrm{D}E(\boldsymbol{\varphi}), J\circ(\mathrm{D}E(\boldsymbol{\varphi}+\boldsymbol{g}) - \mathrm{D}E(\boldsymbol{\varphi}) - \mathrm{D}^2E(\boldsymbol{\varphi})\boldsymbol{g})\rangle \\
&= \langle (\Delta + f'(\varphi))\psi, g\rangle - \langle \dot\psi, \dot g\rangle - \langle \dot\varphi, f(\varphi+g) - f(\varphi) - f'(\varphi)g\rangle.
\end{aligned}$$

The first term is $\lesssim C_0\mathrm{e}^{-3\kappa|t|}$, due to (2.19) and (3.10), hence it is negligible (by enlarging C_0 if necessary). Inequality (2.18) implies that the second term can be replaced by $-\frac{b}{\lambda}(\lambda' - b)\langle \Lambda_0\Lambda W_{\underline{\lambda}}, \dot g\rangle$, which in turn can be replaced by $-b(\lambda' - b)\langle A_0(\lambda)\Lambda W_{\underline{\lambda}}, \dot g\rangle$, thanks to (3.42). From (3.43) we infer that the third term can be replaced by $b\cdot\langle A(\lambda)\varphi, f(\varphi+g) - f(\varphi) - f'(\varphi)g\rangle$. Using formula (3.40) with $h_1 = \varphi$ and $h_2 = g$ we obtain

$$I'(t) \simeq -b(\lambda' - b)\cdot\langle A_0(\lambda)\Lambda W_{\underline{\lambda}}, \dot g\rangle - b\cdot\langle A(\lambda)g, f(\varphi+g) - f(\varphi)\rangle. \tag{3.55}$$

Step 2 (Derivative of the virial functional) We compute:

$$\begin{aligned}
(bJ)'(t) &= b'\int \dot g\cdot A_0(\lambda)g\,\mathrm{d}x + b\lambda'\int \dot g\cdot\partial_\lambda A_0(\lambda)g\,\mathrm{d}x \\
&\quad + b\int \dot g\cdot A_0(\lambda)(\dot g - \psi)\,\mathrm{d}x + b\int(\Delta g + f(\varphi+g) - f(\varphi) - \dot\psi)\cdot A_0(\lambda)g\,\mathrm{d}x.
\end{aligned} \tag{3.56}$$

The first two terms are negligible thanks to Lemma 3.12. Consider the third term on the right in (3.56). An integration by parts yields $\int \dot g \cdot A_0(\lambda)\dot g\,\mathrm{d}x = 0$. Since $A_0(\lambda) : \dot H^1 \to L^2$ is a bounded operator, from (2.17) we see that

$$b\int \dot g \cdot A_0(\lambda)\psi\,\mathrm{d}x \simeq -b(\lambda' - b)\cdot\langle A_0(\lambda)\Lambda W_{\underline{\lambda}}, \dot g\rangle.$$

Consider the forth term on the right in (3.56). The term $b\int \dot\psi \cdot A_0(\lambda) g\,\mathrm{d}x$ is negligible. Using (3.41) and the fact that $A_0(\lambda) = A(\lambda) + \frac{1}{6\lambda}\Delta q(\frac{\cdot}{\lambda})$ we get

$$\begin{aligned} b\int(\Delta g + f(\varphi+g) - f(\varphi))\cdot A_0(\lambda) g\,\mathrm{d}x &\leq b\cdot\langle A(\lambda)g, f(\varphi+g) - f(\varphi)\rangle \\ -\frac{b}{\lambda}\int_{|x|\leq R\lambda}|\nabla g|^2\,\mathrm{d}x &+ \frac{b}{\lambda}\int\frac{1}{6}\Delta q\big(\frac{\cdot}{\lambda}\big)(f(\varphi+g)-f(\varphi))g\,\mathrm{d}x + cC_0^2\mathrm{e}^{-3\kappa|t|}, \end{aligned}$$

with a small constant c. Putting everything back together and using (3.44) we get

$$\begin{aligned} (bJ)'(t) \leq -b(\lambda' - b)\cdot\langle A_0(\lambda)\Lambda W_{\underline{\lambda}}, \dot g\rangle &+ b\cdot\langle A(\lambda)g, f(\varphi+g) - f(\varphi)\rangle \\ + \frac{b}{\lambda}\Big(-\int_{|x|\leq R\lambda}|\nabla g|^2\,\mathrm{d}x &+ \int f'(W_\lambda)g^2\,\mathrm{d}x\Big) + cC_0^2\mathrm{e}^{-3\kappa|t|}. \end{aligned}$$

Step 3 (Localized coercivity) Taking the sum of (3.55) and (3.56) we obtain

$$H'(t) \leq \frac{b}{\lambda}\Big(-\int_{|x|\leq R\lambda}|\nabla g|^2\,\mathrm{d}x + \int f'(W_\lambda)g^2\,\mathrm{d}x\Big) + cC_0^2\mathrm{e}^{-3\kappa|t|}.$$

Recall that $\big|\big\langle\frac{1}{\lambda}\mathcal{Y}_{\underline{\lambda}}, g\big\rangle\big| \lesssim |a_1^+| + |a_1^-|$, hence (3.30) (after rescaling) together with (3.14), (3.11) and (3.18) imply that

$$-\int_{|x|\leq R\lambda}|\nabla g|^2\,\mathrm{d}x + \int f'(W_\lambda)g^2\,\mathrm{d}x \lesssim (cC_0^2 + 1)\mathrm{e}^{-3\kappa|t|},$$

with $c > 0$ as small as we wish (by taking R large enough). Enlarging C_0 if necessary we arrive at (3.51). □

3.6 Shooting argument and passing to a limit

We are ready to give a proof of the main result of the paper, following a well-known scheme introduced in [64] and [56].

Proof of Theorem 1.

Step 1 Let t_n be a decreasing sequence converging to $-\infty$. For n large and

$$a_0 \in \mathcal{A} := \big[-\frac{1}{2}\mathrm{e}^{-\frac{3}{2}\kappa|T_n|}, \frac{1}{2}\mathrm{e}^{-\frac{3}{2}\kappa|T_n|}\big],$$

let $\boldsymbol{u}_n^{a_0}(t) : [t_n, T_+) \to \mathcal{E}$ denote the solution of (1.1) with initial data (3.5). We will prove that there exists a_0 such that $T_+ > T_0$ and for $\boldsymbol{u} = \boldsymbol{u}_n^{a_0}$ inequalities (3.12), (3.13) hold for $t \in [t_n, T_0]$.

Suppose that this is not the case. For each $a_0 \in \mathcal{A}$, let $T_1 = T_1(a_0)$ be the last time such that (3.12) and (3.13) hold for $t \in [t_n, T_1)$. Since $\{\boldsymbol{\varphi}(t) : t \in [t_n, T_1]\}$ is a compact set, Corollary A.3 implies that $T_+ > T_1$. Suppose that $|a_1^+(T_1)| < \mathrm{e}^{-\frac{3}{2}\kappa|T_1|}$. Then Proposition 3.3 would imply that (3.12) and (3.13) hold on some neighborhood of T_1, contradicting its definition. Thus $a_1^+(T_1) = \mathrm{e}^{-\frac{3}{2}\kappa|T_1|}$ or $a_1^+(T_1) = -\mathrm{e}^{-\frac{3}{2}\kappa|T_1|}$. Let $\mathcal{A}_+ \subset \mathcal{A}$ be the set of $a_0 \in \mathcal{A}$ which lead to $a_1^+(T_1) = \mathrm{e}^{-\frac{3}{2}\kappa|T_1|}$ and let $\mathcal{A}_- \subset \mathcal{A}$ be the set of $a_0 \in \mathcal{A}$ which lead to $a_1^+(T_1) = -\mathrm{e}^{-\frac{3}{2}\kappa|T_1|}$. We have proved that $\mathcal{A} = \mathcal{A}_+ \cup \mathcal{A}_-$. We will show that $\mathcal{A}_+$ and $\mathcal{A}_-$ are open sets, which will lead to a contradiction since $\mathcal{A}$ is connected.

Let $a_0 \in \mathcal{A}_+$. This implies that there exists the first T_2 such that $a_1^+(T_2) > \frac{3}{4}\mathrm{e}^{-\frac{3}{2}\kappa|T_2|}$. Hence for a solution $\widetilde{\boldsymbol{u}}(t)$ corresponding to $\widetilde{a}_0$ close to a_0 we will have (by continuity of the flow) $\widetilde{a}_1^+(T_2) > \frac{3}{4}\mathrm{e}^{-\frac{3}{2}\kappa|T_2|}$ and $|\widetilde{a}_1^+(t)| < \mathrm{e}^{-\frac{3}{2}\kappa|t|}$ for $t \in [t_n, T_2]$. Suppose that $\widetilde{a}_0 \in \mathcal{A}_-$. Hence there exists the first $T_3 > T_2$ such that $\widetilde{a}_1^+(T_3) = \frac{3}{4}\mathrm{e}^{-\frac{3}{2}\kappa|T_3|}$. Estimate (3.17) yields $\widetilde{a}_1^+(T_3) \gtrsim \mathrm{e}^{-\frac{1}{2}\kappa|T_3|} \gg \mathrm{e}^{-\frac{3}{2}\kappa|T_3|}$, which is a contradiction. Hence $\mathcal{A}_+$ is open and analogously $\mathcal{A}_-$ is open.

Step 2 Call $\boldsymbol{u}_n$ the solution found in Step 1. From (3.12), (3.13) and (2.11) we deduce that there exists a constant $C_1 > 0$ independent of n such that

$$\|\boldsymbol{u}_n(t) - (W_{\frac{1}{\kappa}\mathrm{e}^{-\kappa|t|}} - W, -\mathrm{e}^{-\kappa|t|}\Lambda W_{\underline{\frac{1}{\kappa}\mathrm{e}^{-\kappa|t|}}})\|_{\mathcal{E}} \leq C_1 \cdot \mathrm{e}^{-\frac{3}{2}\kappa|t|}, \qquad \text{for } t \in [t_n, T_0]. \tag{3.57}$$

The sequence $\boldsymbol{u}_n(T_0)$ is bounded in $\mathcal{E}$, hence its subsequence (which we still denote $\boldsymbol{u}_n$) has a weak limit $\boldsymbol{u}_0$. Let $\boldsymbol{u}(t)$ we the solution of (1.1) with the initial data $\boldsymbol{u}(T_0) = \boldsymbol{u}_0$. Let $T < T_0$. In view of (3.57), for large n the sequence $\boldsymbol{u}_n$ satisfies the assumptions of Corollary A.4 on the time interval $[T, T_0]$, hence $\boldsymbol{u}_n(T) \rightharpoonup \boldsymbol{u}(T)$. Passing to the weak limit in (3.57) finishes the proof. □

Remark 3.14. Note that only the instability component $a_1^+(t)$ is treated via a topological argument, whereas $a_2^+(t)$ is estimated directly. This depends heavily on the (incidental) fact that the bootstrap bound $\mathrm{e}^{-\frac{3}{2}\kappa|t|}$ is asymptotically smaller than $\mathrm{e}^{-\nu|t|}$. Were it not the case, we would have to use a topological argument based on the Brouwer fixed point theorem, as in the work of Côte, Martel and Merle [18].

4 Bubble-antibubble for the radial critical Yang-Mills equation

4.1 Notation

In this section we denote $\|v\|^2_{L^2(r_1\leq r\leq r_2)} := 2\pi\int_{r_1}^{r_2}|v(r)|^2\,r\mathrm{d}r$. If r_1 or r_2 is not precised, then it should be understood that $r_1 = 0$, resp. $r_2 = +\infty$. The corresponding scalar product is denoted $\langle v, w\rangle := 2\pi\int_0^{+\infty} v(r)\cdot w(r)\,r\mathrm{d}r$.

Recall that $\|v\|^2_{\mathcal{H}} := 2\pi\int_0^{+\infty}\big(|\partial_r v(r)|^2 + |\frac{2}{r}v(r)|^2\big)\,r\mathrm{d}r$. A change of variables shows that $v(r) \in \mathcal{H} \Leftrightarrow v(\mathrm{e}^x) \in H^1(\mathbb{R})$, in particular $\|v\|_{L^\infty} \lesssim \|v\|_{\mathcal{H}}$ (this change of variables, very helpful in proving coercivity lemmas, can be found in [35]). Another useful way of understanding the space $\mathcal{H}$ is to consider the transformation $\widetilde{v}(\mathrm{e}^{i\theta}r) := \mathrm{e}^{2i\theta}v(r)$, which is an isometric embedding of $\mathcal{H}$ in $\dot{H}^1(\mathbb{R}^2;\mathbb{R}^2)$, whose image is given by 2-equivariant functions in $\dot{H}^1(\mathbb{R}^2;\mathbb{R}^2)$. Let $\mathcal{H}^*$ be the dual space of $\mathcal{H}$ for the pairing $\langle\cdot,\cdot\rangle$. The embedding just described identifies $\mathcal{H}^*$ with the 2-equivariant distributions in $\dot{H}^{-1}(\mathbb{R}^2;\mathbb{R}^2)$.

We denote $X^0 := L^2 \cap \mathcal{H}$ and $X^1 := \{v \in \mathcal{H} : \partial_r v \in \mathcal{H} \text{ and } \frac{1}{r}v \in \mathcal{H}\}$. The generators of the $\mathcal{H}$-critical and the L^2-critical scale change will be denoted respectively $\Lambda := r\partial_r$ and $\Lambda_0 := 1 + r\partial_r$.

4.2 Linearized equation and formal computation

Linearizing $-\partial_r^2 u - \frac{1}{r}\partial_r u + \frac{4}{r^2}u(1-u)(1-\frac{1}{2}u)$ around $u = W$ we obtain the operator

$$L := -\partial_r^2 - \frac{1}{r}\partial_r - \frac{1}{r^2}(2 - 6(W(r)-1)^2) = -\partial_r^2 - \frac{1}{r}\partial_r + \frac{1}{r^2}(4 - 6\Lambda W).$$

We study solutions behaving like $\boldsymbol{u}(t) \simeq -\boldsymbol{W} + \boldsymbol{W}_{\lambda(t)}$ with $\lambda(t) \to 0$ as $t \to -\infty$. As in Subsection 2.2, we expand

$$\boldsymbol{u}(t) = -\boldsymbol{W} + \boldsymbol{U}^{(0)}_{\lambda(t)} + b(t)\cdot \boldsymbol{U}^{(1)}_{\lambda(t)} + b(t)^2 \cdot \boldsymbol{U}^{(2)}_{\lambda(t)},$$

with $b(t) = \lambda'(t)$, $\boldsymbol{U}^{(0)} := (W, 0)$ and $\boldsymbol{U}^{(1)} := (0, -\Lambda W)$. This gives

$$\partial_t^2 u(t) = -b'(t)(\Lambda W)_{\underline{\lambda(t)}} + \frac{b(t)^2}{\lambda(t)}(\Lambda_0 \Lambda W)_{\underline{\lambda(t)}} + \text{lot}. \tag{4.1}$$

Let us restrict our attention to the region $r \leq \sqrt{\lambda(t)}$. We will see that the region $r \geq \sqrt{\lambda(t)}$ will not have much influence on the dynamics of our system. For $r \leq \sqrt{\lambda}$ we have $W \ll W_\lambda$, hence

$$4u(1-u)(1-\frac{1}{2}u) \simeq 4W_\lambda(1-W_\lambda)(1-\frac{1}{2}W_\lambda) + (-4 + 6\Lambda W_\lambda)W + \text{lot}.$$

Since $\partial_r^2 W + \frac{1}{r}\partial_r W \simeq \frac{4}{r^2}W + \text{lot}$ for $r \leq \sqrt{\lambda}$, we get

$$\partial_r^2 u + \frac{1}{r}\partial_r u - \frac{4}{r^2}u(1-u)(1-\frac{1}{2}u) = -\frac{b^2}{\lambda}(LU^{(2)})_{\underline{\lambda}} - \frac{1}{r^2}6W_\lambda \cdot W + \text{lot}.$$

We can further simplify this using the fact that $W(r) \simeq 2r^2$:

$$\partial_r^2 u + \frac{1}{r}\partial_r u - \frac{4}{r^2}u(1-u)(1-\frac{1}{2}u) = -\frac{b^2}{\lambda}(LU^{(2)})_{\underline{\lambda}} - 12\Lambda W_\lambda + \text{lot},$$

thus (4.1) yields

$$LU^{(2)} = -\Lambda_0 \Lambda W + \frac{\lambda}{b^2}(b' - 12\lambda)\Lambda W.$$

As in Subsection 2.2, we find that the best choice of the formal parameter equations is

$$\lambda'(t) = b(t), \qquad b'(t) = \kappa^2 \lambda(t), \qquad \text{with } \kappa := 2\sqrt{3}.$$

Remark 4.1. The main term of the interaction is *exactly* cancelled by the term $-b'\Lambda W_{\underline{\lambda}}$ for our choice of the parameters. We have seen that it is not the case for the power nonlinearity and in the next section we will see that it is not the case either for the critical equivariant wave map equation.

4.3 Bounds on the error of the ansatz

Fix $\mathcal{Z} \in C_0^\infty((0, +\infty))$ such that

$$\int_0^{+\infty} \mathcal{Z}(r)\cdot \Lambda W(r)\frac{\mathrm{d}r}{r} > 0. \tag{4.2}$$

By a direct computation we find $L\big(\frac{r^4}{(1+r^2)^2}\big) = \frac{r^2(3-r^2)}{(1+r^2)^3} = -\Lambda_0\Lambda W$. Adding a suitable multiple of $\Lambda W(r)$, we obtain a rational function $Q(r)$ such that

$$LQ = -\Lambda_0\Lambda W, \qquad \int \mathcal{Z}(r)\cdot Q(r)\frac{\mathrm{d}r}{r} = 0, \qquad Q(r)\sim r^2 \text{ as } r\to 0,\ Q(r)\sim 1 \text{ as } r\to+\infty. \tag{4.3}$$

For λ, μ satisfying (2.6) and (2.7) (naturally with $\kappa = 2\sqrt{3}$) we define the approximate solution by the formula

$$\begin{aligned}\varphi(t) &:= -W_{\mu(t)} + W_{\lambda(t)} + S(t),\\ \dot\varphi(t) &:= -b(t)\Lambda W_{\underline{\lambda(t)}},\end{aligned}$$

where

$$\begin{aligned} b(t) &:= \mathrm{e}^{-\kappa|T|} + \kappa^2\int_T^t \frac{\lambda(\tau)}{\mu(\tau)^2}\,\mathrm{d}\tau, && \text{for } t\in[T,T_0],\\ S(t) &:= \chi\cdot b(t)^2 Q_{\lambda(t)}, && \text{for } t\in[T,T_0].\end{aligned}$$

From (4.3) we obtain

$$\|\chi\cdot Q_{\underline{\lambda}}\|_{\mathcal{H}} \lesssim \Big(\int_0^{2/\lambda} 1\cdot r\mathrm{d}r\Big)^{\frac12} + \frac{1}{\lambda}\cdot\Big(\int_0^{2/\lambda}((1+r)^{-1})^2\, r\mathrm{d}r\Big)^{\frac12} \lesssim \sqrt{|t|}\cdot \mathrm{e}^{\kappa|t|},$$

which implies that

$$\|S(t)\|_{\mathcal{H}} \ll \mathrm{e}^{-\frac32\kappa|t|}.$$

Since $\mathcal{Z}$ has compact support, (2.3) implies, for sufficiently small λ,

$$\int \mathcal{Z}_{\underline{\lambda}}\cdot S(t)\frac{\mathrm{d}r}{r} = 0.$$

We denote $f(u) := -4u(1-u)(1-\frac12 u)$ and

$$\begin{aligned}\boldsymbol{\psi}(t) = (\psi(t),\dot\psi(t)) &:= \partial_t\boldsymbol{\varphi}(t) - \mathrm{D}E(\boldsymbol{\varphi}(t))\\ &= \big(\partial_t\varphi(t) - \dot\varphi(t), \partial_t\dot\varphi(t) - (\partial_r^2\varphi(t) + \frac1r\partial_r\varphi(t) + \frac{1}{r^2}f(\varphi(t)))\big).\end{aligned}$$

We have $f'(u) = 2 - 6(u-1)^2$.

Remark 4.2. By a direct computation, $f'(W) = -4 + 6\Lambda W$. Thus the potential term of the linearized operator contains a non-localized part -4 and a localized part $6\Lambda W$. These terms need to be treated in different ways. This is a known issue coming from the fact that $f(u)$ is not really the nonlinearity, as it "hides" the linear part near the stable equilibria: $f(u)\simeq -4u$ near $u=0$ and $f(u)\simeq 4(2-u)$ near $u=2$. Sometimes it is convenient to subtract the linear part from f, but here we work simultaneously near $u=0$ and near $u=2$, so probably it will be simpler to keep f as it is. A similar remark could be made in the case of the equivariant wave map equation.

Lemma 4.3. *Suppose that for $t\in[T,T_0]$ there holds $|\lambda'(t)|\lesssim \mathrm{e}^{-\kappa|t|}$ and $|\mu'(t)|\lesssim \mathrm{e}^{-\kappa|t|}$. Then*

$$\|\psi(t) - \mu'(t)\frac{1}{\mu(t)}\Lambda W_{\mu(t)} + (\lambda'(t)-b(t))\frac{1}{\lambda(t)}\Lambda W_{\lambda(t)}\|_{\mathcal{H}} \lesssim \mathrm{e}^{-\frac32\kappa|t|}, \tag{4.4}$$

$$\|\dot\psi(t) - \frac{b(t)}{\lambda(t)}(\lambda'(t)-b(t))\Lambda_0\Lambda W_{\underline{\lambda(t)}}\|_{L^2} \lesssim \mathrm{e}^{-\frac32\kappa|t|}, \tag{4.5}$$

$$\|(-\partial_r^2 - \frac1r\partial_r - \frac{1}{r^2}f'(\varphi(t)))\psi(t)\|_{\mathcal{H}^*} \lesssim \mathrm{e}^{-\frac32\kappa|t|}. \tag{4.6}$$

Proof. The proof of (4.4) is the same as the proof of (2.17).

In order to prove (4.5), we treat separately the regions $r \leq \sqrt{\lambda}$ and $r \geq \sqrt{\lambda}$. We will show that

$$\left\|\frac{1}{r^2}\big(f(\varphi) - f(W_\lambda) + 2\frac{r^2}{\mu^2}f'(W_\lambda) - b^2 f'(W_\lambda)Q_\lambda\big)\right\|_{L^2(r\leq\sqrt{\lambda})} \lesssim \mathrm{e}^{-\frac{3}{2}\kappa|t|}. \tag{4.7}$$

Since f is a polynomial of degree 3, we have

$$f(\varphi) = f(W_\lambda) + f'(W_\lambda)(-W_\mu + S) + \frac{1}{2}f''(W_\lambda)(-W_\mu + S)^2 + \frac{1}{6}f'''(W_\lambda)(-W_\mu + S)^3.$$

We treat all the terms one by one. We have $\big|W_\mu(r) - 2\frac{r^2}{\mu^2}\big| \lesssim r^4$. Since $|f'(W_\lambda)| \lesssim 1$, this implies

$$\left\|\frac{1}{r^2}f'(W_\lambda)(W_\mu - 2r^2)\right\|_{L^2} \lesssim \Big(\int_0^{\sqrt{\lambda}} r^4\cdot r\mathrm{d}r\Big)^{\frac{1}{2}} \lesssim \mathrm{e}^{-\frac{3}{2}\kappa|t|}.$$

Since r is small, we have $S(r) = b^2 Q_\lambda(r)$, and the corresponding term is subtracted in (4.7).

Next, notice that $|W_\mu| + |S| \lesssim r^2$. Since $|f''(W_\lambda)| \lesssim 1$, this implies

$$\left\|\frac{1}{r^2}f''(W_\lambda)(-W_\mu + S)^2\right\|_{L^2} \lesssim \Big(\int_0^{\sqrt{\lambda}} \big(\frac{1}{r^2}\cdot r^4\big)^2 r\mathrm{d}r\Big)^{\frac{1}{2}} \sim \lambda^{\frac{3}{2}} \lesssim \mathrm{e}^{-\frac{3}{2}\kappa|t|}.$$

The last term is estimated in a similar way. This finishes the proof of (4.7).

By a direct computation $\big|\big(\partial_r^2 + \frac{1}{r}\partial_r\big)W_\mu - \frac{8}{\mu^2}\big| \lesssim r^2$, hence

$$\left\|\big(\partial_r^2 + \frac{1}{r}\partial_r\big)W_\mu - \frac{8}{\mu^2}\right\|_{L^2(r\leq\sqrt{\lambda})} \lesssim \mathrm{e}^{-\frac{3}{2}\kappa|t|}. \tag{4.8}$$

From (4.7), (4.8) and the definition of $\varphi(t)$ we have

$$\left\|\big((\partial_r^2 + \frac{1}{r}\partial_r)\varphi + \frac{1}{r^2}f(\varphi)\big) - \big(-\frac{2}{\mu^2}f'(W_\lambda) - \frac{b^2}{\lambda}(LQ)_{\underline{\lambda}} - \frac{8}{\mu^2}\big)\right\|_{L^2(r\leq\sqrt{\lambda})} \lesssim \mathrm{e}^{-\frac{3}{2}\kappa|t|}.$$

Using (4.3) and the relation $4 + f'(W) = 6\Lambda W$ we can rewrite this as

$$\left\|\big((\partial_r^2 + \frac{1}{r}\partial_r)\varphi + \frac{1}{r^2}f(\varphi)\big) - \big(-\frac{12}{\mu^2}\Lambda W_\lambda + \frac{b^2}{\lambda}\Lambda_0\Lambda W_{\underline{\lambda}}\big)\right\|_{L^2(r\leq\sqrt{\lambda})} \lesssim \mathrm{e}^{-\frac{3}{2}\kappa|t|},$$

which is equivalent to (4.5), restricted to the region $r \leq \sqrt{\lambda}$.

Consider the region $r \geq \sqrt{\lambda}$. Developping f at $2 - W_\mu$ we get

$$f(\varphi) = f(2-W_\mu)+f'(2-W_\mu)(W_\lambda-2+S)+\frac{1}{2}f''(2-W_\mu)(W_\lambda-2+S)^2+\frac{1}{6}f'''(2-W_\mu)(W_\lambda-2+S)^3.$$

From this and the relations $f(2-W_\mu) = -f(W_\mu)$, $f'(2-W_\mu) = f'(W_\mu)$, we obtain a pointwise bound

$$|f(\varphi) + f(W_\mu) + f'(W_\mu)(2 - W_\lambda)| \lesssim |S| + |2 - W_\lambda|^2. \tag{4.9}$$

Since $|S(r)| \lesssim b^2$, we have

$$\left\|\frac{1}{r^2}S\right\|_{L^2(r\geq\sqrt{\lambda})} \lesssim b^2\Big(\int_{\sqrt{\lambda}}^{+\infty} r^{-4}\, r\mathrm{d}r\Big)^{\frac{1}{2}} \lesssim \mathrm{e}^{-\frac{3}{2}\kappa|t|}. \tag{4.10}$$

There holds $|2 - W_\lambda(r)|^2 \lesssim \frac{\lambda^4}{r^4}$, hence

$$\left\|\frac{1}{r^2}|2 - W_\lambda(r)|^2\right\|_{L^2(r\geq\sqrt{\lambda})} \lesssim \lambda^4\Big(\int_{\sqrt{\lambda}}^{+\infty} r^{-12}\, r\mathrm{d}r\Big)^{\frac{1}{2}} \lesssim \mathrm{e}^{-\frac{3}{2}\kappa|t|}. \tag{4.11}$$

Since $|W_\mu - 2| \lesssim r^2$, there holds $|f'(W_\mu) + 4| \lesssim r^2$. We also have $|2 - W_\lambda| \lesssim \frac{\lambda^2}{r^2}$ and $|(2 - W_\lambda) - \frac{2\lambda^2}{r^2}| \lesssim \frac{\lambda^4}{r^4}$, hence

$$\begin{aligned}&\Big\|\frac{1}{r^2}f'(W_\mu)(2 - W_\lambda) + \frac{8\lambda^2}{r^4}\Big\|_{L^2(r\geq\sqrt{\lambda})} \lesssim \Big\|\frac{\lambda^2}{r^2} + \frac{\lambda^4}{r^6}\Big\|_{L^2(r\geq\sqrt{\lambda})} \\ &\lesssim \lambda^2\Big(\int_{\sqrt{\lambda}}^{+\infty} r^{-4}\,r\mathrm{d}r\Big)^{\frac12} + \lambda^4\Big(\int_{\sqrt{\lambda}}^{+\infty} r^{-12}\,r\mathrm{d}r\Big)^{\frac12} \lesssim \mathrm{e}^{-\frac32|\kappa|t}.\end{aligned} \tag{4.12}$$

Inserting (4.10), (4.11) and (4.12) into (4.9) we obtain

$$\Big\|f(\varphi) + f(W_\mu) - \frac{8\lambda^2}{r^4}\Big\|_{L^2(r\geq\sqrt{\lambda})} \lesssim \mathrm{e}^{-\frac32\kappa|t|}. \tag{4.13}$$

A direct computation shows that $(\partial_r^2 + \frac1r\partial_r)W_\lambda(r) = -\frac{8\lambda^2}{r^4} + O(\frac{\lambda^4}{r^6} + \frac{\lambda^8}{r^{10}})$, hence

$$\Big\|(\partial_r^2 + \frac1r\partial_r)W_\lambda + \frac{8\lambda^2}{r^4}\Big\|_{L^2(r\geq\sqrt{\lambda})} \lesssim \mathrm{e}^{-\frac32\kappa|t|}. \tag{4.14}$$

We have $\partial_r S = \chi' b^2\cdot Q_\lambda + \chi b^2\cdot(Q_\lambda)'$ and $\partial_r^2 S = \chi'' b^2\cdot Q_\lambda + 2\chi' b^2\cdot(Q_\lambda)' + \chi b^2\cdot(Q_\lambda)''$. There holds $|Q| \lesssim 1$, $|Q'| \lesssim \frac1r$ and $|Q''| \lesssim \frac{1}{r^2}$, which implies $|Q_\lambda| \lesssim 1$, $|(Q_\lambda)'| \lesssim \frac1r$ and $|(Q_\lambda)''| \lesssim \frac{1}{r^2}$. This gives

$$\begin{aligned}\|\chi b^2\cdot(Q_\lambda)''\|_{L^2(r\geq\sqrt{\lambda})} + \Big\|\frac1r\chi b^2\cdot(Q_\lambda)'\Big\|_{L^2(r\geq\sqrt{\lambda})} &\lesssim b^2\Big(\int_{\sqrt{\lambda}}^{+\infty} r^{-4}\,r\mathrm{d}r\Big)^{\frac12} \lesssim \frac{b^2}{\sqrt{\lambda}} \lesssim \mathrm{e}^{-\frac32\kappa|t|}, \\ \|\chi' b^2\cdot(Q_\lambda)'\|_{L^2(r\geq\sqrt{\lambda})} + \|\chi'' b^2\cdot Q_\lambda\|_{L^2(r\geq\sqrt{\lambda})} + \Big\|\frac1r\chi' b^2\cdot Q_\lambda\Big\|_{L^2(r\geq\sqrt{\lambda})} &\lesssim b^2 \ll \mathrm{e}^{-\frac32\kappa|t|},\end{aligned} \tag{4.15}$$

hence $\|(\partial_r^2 + \frac1r\partial_r)S\|_{L^2(r\geq\sqrt{\lambda})} \lesssim \mathrm{e}^{-\frac32\kappa|t|}$. Together with (4.13) and (4.14) this proves that

$$\|(\partial_r^2 + \frac1r\partial_r)\varphi + \frac{1}{r^2}f(\varphi)\|_{L^2(r\geq\sqrt{\lambda})} \lesssim \mathrm{e}^{-\frac32\kappa|t|}.$$

Since $\|\Lambda W_{\underline{\lambda}}\|_{L^2(r\geq\sqrt{\lambda})} + \|\Lambda_0\Lambda W_{\underline{\lambda}}\|_{L^2(r\geq\sqrt{\lambda})} \lesssim (\int_{1/\sqrt{\lambda}}^{+\infty}\frac{1}{r^4}\,r\mathrm{d}r)^{\frac12} \lesssim \sqrt{\lambda}$, the other terms appearing in (4.5) are $\lesssim \mathrm{e}^{-\frac32\kappa|t|}$. This finishes the proof of (4.5).

The proof of (4.6) is very similar to the proof of (2.19), hence we will just indicate the differences. Since $\|W_\lambda\|_{L^\infty} + \|W_\mu\|_{L^\infty} \lesssim 1$, we have $|f'(-W_\mu + W_\lambda) - f'(W_\lambda)| \lesssim W_\mu$ and $|f'(-W_\mu + W_\lambda) - f'(W_\mu)| = |f'(W_\mu + (2 - W_\lambda)| - f'(W_\mu)| \lesssim 2 - W_\lambda$. Next, we check that $\|\frac{1}{r^2}W_\mu\cdot\Lambda W_\lambda\|_{L^1} \lesssim \mathrm{e}^{-\frac32\kappa|t|}$ (recall that $\mathcal{H} \subset L^\infty$, hence $L^1(\mathbb{R}^2) \subset \mathcal{H}^*$). To do this, we consider separately $r \leq 1$ and $r \geq 1$:

$$\begin{aligned}\Big\|\frac{1}{r^2}W_\mu\cdot\Lambda W_\lambda\Big\|_{L^1(r\leq1)} &\lesssim \|\Lambda W_\lambda\|_{L^1(r\leq1)} = \lambda^2\|\Lambda W\|_{L^1(r\leq1/\lambda)} \lesssim \lambda^2|\log\lambda| \ll \mathrm{e}^{-\frac32\kappa|t|}, \\ \Big\|\frac{1}{r^2}W_\mu\cdot\Lambda W_\lambda\Big\|_{L^1(r\geq1)} &\lesssim \|\Lambda W_\lambda\|_{L^\infty(r\geq1)} = \big|\Lambda W\big(\frac1\lambda\big)\big| \sim \lambda^2 \ll \mathrm{e}^{-\frac32\kappa|t|}.\end{aligned}$$

Finally, we check that $\|\frac{1}{r^2}(2 - W_\lambda)\cdot\Lambda W_\mu\|_{L^1} \lesssim \mathrm{e}^{-\frac12\kappa|t|}$, again dividing into $r \leq 1$ and $r \geq 1$:

$$\begin{aligned}\|2 - W_\lambda\|_{L^1(r\leq1)} &= \Big\|\frac{1}{1 + (r/\lambda)^2}\Big\|_{L^1(r\leq1)} = \lambda^2\Big\|\frac{1}{1 + r^2}\Big\|_{L^1(r\leq1/\lambda)} \sim \lambda^2|\log\lambda| \ll \mathrm{e}^{-\frac32\kappa|t|}, \\ \|2 - W_\lambda\|_{L^1(r\geq1)} &= \frac{2\lambda^2}{1 + \lambda^2} \ll \mathrm{e}^{-\frac32\kappa|t|}.\end{aligned}$$

This allows to conclude, since $\|f'(\varphi) - f'(-W_\mu + W_\lambda)\|_{L^\infty} \lesssim \|S\|_{L^\infty} \lesssim \mathrm{e}^{-\frac32\kappa|t|}$. □

4.4 Modulation

Having defined the approximate solution $\boldsymbol{\varphi}(t)$, we will now analyse exact solutions close to $\boldsymbol{\varphi}(t)$. The initial data are

$$\boldsymbol{u}(T) = \big(-W + W_{\frac{1}{\kappa}\mathrm{e}^{-\kappa|T|}}, -\mathrm{e}^{-\kappa|T|}\underline{\Lambda W_{\frac{1}{\kappa}\mathrm{e}^{-\kappa|T|}}}\big)$$

(there is no linear instability in the case of the Yang-Mills equation).

Similarly as in Subsection 3.2, we choose the modulation parameters $\lambda(t)$ and $\mu(t)$ which verify

$$\int \mathcal{Z}_{\underline{\mu}} \cdot \big(u(t) - (-W_{\mu(t)} + W_{\lambda(t)})\big)\frac{\mathrm{d}r}{r} = 0, \qquad \int \mathcal{Z}_{\underline{\lambda}} \cdot \big(u(t) - (-W_{\mu(t)} + W_{\lambda(t)})\big)\frac{\mathrm{d}r}{r} = 0.$$

We define $\boldsymbol{g}(t)$ by

$$\boldsymbol{u}(t) = \boldsymbol{\varphi}(t) + \boldsymbol{g}(t).$$

It satisfies, cf. (3.14),

$$\big\langle \frac{1}{\lambda(t)}\mathcal{Z}_{\underline{\lambda(t)}}, g(t)\big\rangle = 0, \qquad \big|\big\langle \frac{1}{\mu(t)}\mathcal{Z}_{\underline{\mu(t)}}, g(t)\big\rangle\big| \leq c\cdot \mathrm{e}^{-\frac{3}{2}\kappa|t|}.$$

The functions $\lambda(t)$ and $\mu(t)$ are C^1 and

$$|\lambda'(t) - b(t)| + |\mu'(t)| \lesssim \|\boldsymbol{g}(t)\|_{\mathcal{E}}^2 + c\cdot \mathrm{e}^{-\frac{3}{2}\kappa|t|}$$

with $c > 0$ arbitrarily small, cf. (3.15).

4.5 Coercivity

Recall that $f'(W) = -4 + 6\Lambda W$. In the next lemma, it is useful to separate these two terms, see Remark 4.2.

Lemma 4.4. *There exist constants $c, C > 0$ such that*

- *for all $g \in \mathcal{H}$ there holds*

$$\begin{aligned}&\int_0^{+\infty}\big((g')^2 + \frac{4}{r^2}g^2\big)\,r\mathrm{d}r - \int_0^{+\infty}\frac{6}{r^2}\Lambda W g^2\,r\mathrm{d}r \\ &\geq c\int_0^{+\infty}\big((g')^2 + \frac{4}{r^2}g^2\big)\,r\mathrm{d}r - C\Big(\int_0^{+\infty}\mathcal{Z}\cdot g\frac{\mathrm{d}r}{r}\Big)^2,\end{aligned} \tag{4.16}$$

- *if $r_1 > 0$ is large enough, then for all $g \in \mathcal{H}$ there holds*

$$\begin{aligned}&(1-2c)\int_0^{r_1}\big((g')^2 + \frac{4}{r^2}g^2\big)\,r\mathrm{d}r + c\int_{r_1}^{+\infty}\big((g')^2 + \frac{4}{r^2}g^2\big)\,r\mathrm{d}r \\ &- \int_0^{+\infty}\frac{6}{r^2}\Lambda W g^2\,r\mathrm{d}r \geq -C\Big(\int_0^{+\infty}\mathcal{Z}\cdot g\frac{\mathrm{d}r}{r}\Big)^2,\end{aligned} \tag{4.17}$$

- *if $r_2 > 0$ is small enough, then for all $g \in \mathcal{H}$ there holds*

$$\begin{aligned}&(1-2c)\int_{r_2}^{+\infty}\big((g')^2 + \frac{4}{r^2}g^2\big)\,r\mathrm{d}r + c\int_0^{r_2}\big((g')^2 + \frac{4}{r^2}g^2\big)\,r\mathrm{d}r \\ &- \int_0^{+\infty}\frac{6}{r^2}\Lambda W g^2\,r\mathrm{d}r \geq -C\Big(\int_0^{+\infty}\mathcal{Z}\cdot g\frac{\mathrm{d}r}{r}\Big)^2,\end{aligned} \tag{4.18}$$

Proof. Let $\widetilde g(x) := g(\mathrm{e}^x)$ and $\widetilde{\mathcal Z}(x) := \mathcal Z(\mathrm{e}^x)$. One computes that $f'(W(\mathrm{e}^x)) = -4+6\,\mathrm{sech}^2(x)$, hence (4.16) is equivalent to

$$\int_{\mathbb R}(\widetilde g')^2 + (4-6\,\mathrm{sech}^2)\widetilde g^2\,\mathrm{d}x \geq c\int_{\mathbb R}\big((\widetilde g')^2+\widetilde g^2\big)\,\mathrm{d}x - C\Big(\int_{\mathbb R}\widetilde{\mathcal Z}\cdot\widetilde g\,\mathrm{d}x\Big)^2. \tag{4.19}$$

This quadratic form corresponds to the classical operator $-\frac{\mathrm{d}^2}{\mathrm{d}x^2}+(4-6\,\mathrm{sech}^2)$, for which 0 is a simple discrete eigenvalue, with the eigenspace spanned by sech^2. Decompose $\widetilde g = a\,\mathrm{sech}^2+g_1$, with $\int\mathrm{sech}^2\cdot g_1\,\mathrm{d}x = 0$. From the Sturm-Liouville theory we obtain

$$\int_{\mathbb R}(\widetilde g')^2 + (4-6\,\mathrm{sech}^2)\widetilde g^2\,\mathrm{d}x = \int_{\mathbb R}(g_1')^2+(4-6\,\mathrm{sech}^2)g_1^2\,\mathrm{d}x \gtrsim \int g_1^2\,r\mathrm{d}r.$$

Let $\mathrm{sech}^2 = b\widetilde{\mathcal Z}+(\mathrm{sech}^2)^\perp$ with $\int\widetilde{\mathcal Z}\cdot(\mathrm{sech}^2)^\perp\,\mathrm{d}x = 0$. Since $\int\widetilde{\mathcal Z}\cdot\mathrm{sech}^2\,\mathrm{d}x>0$, see (4.2), we have $\int_{\mathbb R}\big((\mathrm{sech}^2)^\perp\big)^2\,\mathrm{d}x < \int_{\mathbb R}(\mathrm{sech}^2)^2\,\mathrm{d}x$, hence

$$\begin{aligned}\int g_1^2\,\mathrm{d}x &= a^2\int(\mathrm{sech}^2)^2\,\mathrm{d}x - 2a\int\mathrm{sech}^2\cdot\widetilde g\,\mathrm{d}x + \int\widetilde g^2\,\mathrm{d}x\\ &= a^2\int(\mathrm{sech}^2)^2\,\mathrm{d}x - 2ab\int\widetilde{\mathcal Z}\cdot\widetilde g\,\mathrm{d}x - 2a\int(\mathrm{sech}^2)^\perp\cdot\widetilde g\,\mathrm{d}x + \int\widetilde g^2\,\mathrm{d}x\\ &\geq c\int\widetilde g^2\,\mathrm{d}x - C\Big(\int_{\mathbb R}\widetilde{\mathcal Z}\cdot\widetilde g\,\mathrm{d}x\Big)^2,\end{aligned}$$

which implies (4.19).

With the same change of variable, (4.17) and (4.18) will follow once we prove that

$$\begin{aligned}&(1-2c)\int_{|x|\leq R}\big((g')^2+4g^2\big)\,\mathrm{d}x + c\int_{|x|\geq R}\big((g')^2+4g^2\big)\,\mathrm{d}x\\ &-\int_{\mathbb R}6\,\mathrm{sech}^2 g^2\,\mathrm{d}x \geq -C\Big(\int_{\mathbb R}\widetilde{\mathcal Z}\cdot g\,\mathrm{d}x\Big)^2,\end{aligned} \tag{4.20}$$

provided that R is large enough. To this end, take $\widetilde\chi(x) := \chi\big(\frac{2x}{R}\big)$ and $\widetilde h := \widetilde\chi\cdot\widetilde g$. Since $\widetilde{\mathcal Z}$ has compact support and R is large, we have $\int_{\mathbb R}\widetilde{\mathcal Z}\cdot\widetilde h\,\mathrm{d}x = \int_{\mathbb R}\widetilde{\mathcal Z}\cdot\widetilde g\,\mathrm{d}x$. By a standard integration by parts we get $\int_{\mathbb R}(\widetilde h')^2\,\mathrm{d}x = \int_{\mathbb R}\widetilde\chi^2(\widetilde g')^2\,\mathrm{d}x + \int_{\mathbb R}\frac12\big((\widetilde\chi')^2-\widetilde\chi\widetilde\chi''\big)\widetilde g^2\,\mathrm{d}x$. We notice that $|(\widetilde\chi')^2-\widetilde\chi\widetilde\chi''|\lesssim R^{-2}$ is small, in particular for any $c>0$ there holds $\int_{\mathbb R}\chi^2(\widetilde g')^2\,\mathrm{d}x \geq \int_{\mathbb R}(\widetilde h')^2\,\mathrm{d}x - \frac c2\int\big((\widetilde g')^2+4\widetilde g^2\big)x\,\mathrm{d}x$, if R is large enough. Applying (4.19) with $\widetilde h$ instead of $\widetilde g$ and $3c$ instead of c we obtain

$$\begin{aligned}&(1-3c)\int_{|x|\leq R}\big((\widetilde g')^2+4\widetilde g^2\big)\,\mathrm{d}x - \int_{\mathbb R}6\,\mathrm{sech}^2\,\widetilde\chi^2\widetilde g^2\,\mathrm{d}x\\ &\geq (1-3c)\int_{\mathbb R}\chi^2\big((\widetilde g')^2+4\widetilde g^2\big)\,\mathrm{d}x - \int_{\mathbb R}6\,\mathrm{sech}^2\,\widetilde\chi^2\widetilde g^2\,\mathrm{d}x\\ &\geq (1-3c)\int_{\mathbb R}\big((\widetilde h')^2+4\widetilde h^2\big)\,\mathrm{d}x - \frac c2\int_{\mathbb R}\big((\widetilde g')^2+4\widetilde g^2\big)\,\mathrm{d}x - \int_{\mathbb R}6\,\mathrm{sech}^2\,\widetilde h^2\,\mathrm{d}x\\ &\geq -\frac c2\int_{\mathbb R}\big((\widetilde g')^2+4\widetilde g^2\big)\,\mathrm{d}x - C\Big(\int\widetilde{\mathcal Z}\cdot\widetilde h\,\mathrm{d}x\Big)^2 \geq -\frac c2\int_{\mathbb R}\big((\widetilde g')^2+4\widetilde g^2\big)\,\mathrm{d}x - C\Big(\int_{\mathbb R}\widetilde{\mathcal Z}\cdot\widetilde g\,\mathrm{d}x\Big)^2.\end{aligned}$$

But $6\,\mathrm{sech}^2 \leq 6\,\mathrm{sech}^2\,\widetilde\chi^2+2c$ if R is large enough, and (4.20) follows. □

Lemma 4.5. *There exists a constant $\eta>0$ such that if $\frac{\lambda}{\mu}<\eta$ and $\|\boldsymbol U-(-\boldsymbol W_\mu+\boldsymbol W_\lambda)\|_{\mathcal E}<\eta$, then for all $\boldsymbol g\in\mathcal E$ there holds*

$$\langle\mathrm D^2E(\boldsymbol U)\boldsymbol g,\boldsymbol g\rangle + \Big(\int\frac1\lambda\mathcal Z_{\underline\lambda}\cdot g\frac{\mathrm{d}r}{r}\Big)^2 + \Big(\int\frac1\mu\mathcal Z_{\underline\mu}\cdot g\frac{\mathrm{d}r}{r}\Big)^2 \gtrsim \|\boldsymbol g\|_{\mathcal E}^2.$$

□

The proof is a modification of the proof of Lemma 3.9 and will be skipped.

4.6 Definition of the mixed energy-virial functional

Lemma 4.6. *For any $c > 0$ and $R > 0$ there exists a function $q(r) = q_{c,R}(r) \in C^{3,1}((0,+\infty))$ with the following properties:*

(P1) $q(r) = \frac{1}{2}r^2$ *for* $r \leq R$*,*

(P2) there exists $\widetilde{R} > 0$ *(depending on* c *and* R*) such that* $q(r) \equiv \mathrm{const}$ *for* $r \geq \widetilde{R}$*,*

(P3) $|q'(r)| \lesssim r$ *and* $|q''(r)| \lesssim 1$ *for all* $r > 0$*, with constants independent of* c *and* R*,*

(P4) $q''(r) \geq -c$ *and* $\frac{1}{r}q'(r) \geq -c$*, for all* $r > 0$*,*

(P5) $(\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r})^2 q(r) \leq c \cdot r^{-2}$*, for all* $r > 0$*,*

(P6) $\big|r\big(\frac{q'(r)}{r}\big)'\big| \leq c$*, for all* $r > 0$*.*

Proof. It suffices to prove the result for $R = 1$ since the function $q_R(r) := R^2 q(\frac{r}{R})$ satisfies the listed properties if and only if $q(r)$ does.

First we define $q_0(r)$ by the formula

$$q_0(r) := \begin{cases} \frac{1}{2}\cdot r^2 & r \leq 1 \\ \frac{1}{2}r^2 + c_1\big(\frac{1}{2}(r-1)^2 - \log(r)\frac{r^2-1}{4}\big) & r \geq 1, \end{cases}$$

with c_1 small. A direct computation shows that for $r > 1$ we have

$$\begin{aligned} q_0'(r) &= r\big(1 - \frac{c_1 \log r}{2}\big) + c_1\big(\frac{3}{4}r - 1 + \frac{1}{4r}\big), \\ q_0''(r) &= \big(1 - \frac{c_1\log r}{2}\big) + c_1\big(\frac{1}{4} - \frac{1}{4r^2}\big), \\ q_0'''(r) &= -c_1\frac{r^2-1}{2r^3}, \\ \Big(\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\Big)^2 q_0(r) &= -\frac{c_1}{r^3}. \end{aligned} \tag{4.21}$$

In particular $q_0(1) = \frac{1}{2}$, $q_0'(1) = 1$, $q_0''(1) = 1$ and $q_0'''(1) = 0$, hence $q_0 \in C^{3,1}$.

Let $R_0 := \mathrm{e}^{2/c_1}$. From (4.21) it follows that $q_0(r)$ verifies all the listed properties except for (P2) for $r \leq R_0$. Let $e_j(r) := \frac{1}{j!}r^j \cdot \chi(r)$ for $j \in \{1,2,3\}$, where χ is a standard cut-off function. We define

$$q(r) := \begin{cases} q_0(r) & r \leq R_0 \\ q_0(R_0) + \sum_{j=1}^3 q_0^{(j)}(R_0)\cdot R_0^j \cdot e_j(-1 + R_0^{-1}r) & r \geq R_0. \end{cases}$$

We will show that $q(r)$ has all the required properties if c_1 is small enough. It is clear that $q(x) \in C^{3,1}(\mathbb{R}^6)$. Indeed, it follows from (4.21) that $|q_0^{(j)}(R_0)| \lesssim c_1 R_0^{2-j}$ for $j \in \{1,2,3\}$. For $r > R_0$ and $k \in \{1,2,3,4\}$ we have

$$q^{(k)}(r) = \sum_{j=1}^3 q_0^{(j)}(R_0)\cdot R_0^{j-k} e_j^{(k)}(-1 + R_0^{-1}r) \quad \Rightarrow \quad |q^{(k)}(r)| \lesssim c_1 R_0^{2-k}.$$

Since $q(r) \equiv \mathrm{const}$ for $r \geq 3R_0$, we obtain (P1)–(P6). □

We define the operators $A(\lambda)$ and $A_0(\lambda)$ as follows:

$$[A(\lambda)h](r) := q'\big(\frac{r}{\lambda}\big)\cdot h'(r),$$
$$[A_0(\lambda)h](r) := \big(\frac{1}{2\lambda}q''\big(\frac{r}{\lambda}\big) + \frac{1}{2r}q'\big(\frac{r}{\lambda}\big)\big)h(r) + q'\big(\frac{r}{\lambda}\big)\cdot h'(r).$$

Lemma 4.7. *The operators $A(\lambda)$ and $A_0(\lambda)$ have the following properties:*

- *the families $\{A(\lambda) : \lambda > 0\}$, $\{A_0(\lambda) : \lambda > 0\}$, $\{\lambda\partial_\lambda A(\lambda) : \lambda > 0\}$ and $\{\lambda\partial_\lambda A_0(\lambda) : \lambda > 0\}$ are bounded in $\mathscr{L}(\mathcal{H}; L^2)$, with the bound depending on the choice of the function $q(r)$,*
- *for all $\lambda > 0$ and $h_1, h_2 \in X^1$ there holds*

$$\begin{aligned}\big|\big\langle A(\lambda)h_1, \frac{1}{r^2}\big(f(h_1+h_2) - f(h_1) - f'(h_1)h_2\big)\big\rangle + \big\langle A(\lambda)h_2, \frac{1}{r^2}\big(f(h_1+h_2) - f(h_1) + 4h_2\big)\big\rangle\big| \\ \leq \frac{c_0}{\lambda}\big((\|h_1\|_{\mathcal{H}}^2 + 1)\|h_2\|_{\mathcal{H}}^2 + \|h_2\|_{\mathcal{H}}^4\big),\end{aligned} \tag{4.22}$$

with a constant c_0 arbitrarily small,

- *for all $h \in X^1$ there holds*

$$\big\langle A_0(\lambda)h, \big(\partial_r^2 + \frac{1}{r}\partial_r - \frac{4}{r^2}\big)h\big\rangle \leq \frac{c_0}{\lambda}\|h\|_{\mathcal{H}}^2 - \frac{2\pi}{\lambda}\int_0^{R\lambda}\big((\partial_r h)^2 + \frac{4}{r^2}h^2\big)\,r\mathrm{d}r, \tag{4.23}$$

- *assuming (2.7), for any $c_0 > 0$ there holds*

$$\|\Lambda_0\Lambda W_{\underline{\lambda(t)}} - A_0(\lambda(t))\Lambda W_{\lambda(t)}\|_{L^2} \leq c_0, \tag{4.24}$$
$$\|\dot\varphi(t) + b(t)\cdot A(\lambda(t))\varphi(t)\|_{L^\infty} \leq c_0, \tag{4.25}$$
$$\begin{aligned}\Big|\int_0^{+\infty}\frac{1}{2}\big(q''\big(\frac{r}{\lambda}\big) + \frac{\lambda}{r}q'\big(\frac{r}{\lambda}\big)\big)\frac{1}{r^2}\big(f(\varphi+g) - f(\varphi) + 4g\big)g\,r\mathrm{d}r \\ - \int_0^{+\infty}\frac{1}{r^2}\big(f'(W_\lambda) + 4\big)g^2\,r\mathrm{d}r\Big| \leq c_0C_0^2\mathrm{e}^{-3\kappa|t|},\end{aligned} \tag{4.26}$$

provided that the constant R in the definition of $q(r)$ is chosen large enough.

Remark 4.8. The condition $\partial_r h_1, \frac{1}{r}h_1, \partial_r h_2, \frac{1}{r}h_2 \in \mathcal{H}$ is required only to ensure that the left hand side of (3.40) is well defined, but it does not appear on the right hand side of the estimate. Note also that in (4.22), (4.23) and (4.26) we extract the linear part of f, see Remark 4.2.

Proof. The proof of the first point is the same as in Lemma 3.12.

In (4.22), without loss of generality we may assume that $h_1, h_2 \in C_0^\infty((0,+\infty))$ and that $\lambda = 1$. From the definition of $A(\lambda)$ we have

$$\begin{aligned}&\big\langle A(\lambda)h_1, \frac{1}{r^2}\big(f(h_1+h_2) - f(h_1) - f'(h_1)h_2\big)\big\rangle + \big\langle A(\lambda)h_2, \frac{1}{r^2}\big(f(h_1+h_2) - f(h_1) + 4h_2\big)\big\rangle \\ &= \int_0^{+\infty} q'h_1'\cdot\frac{1}{r^2}\big(f(h_1+h_2) - f(h_1) - f'(h_1)\big) + q'h_2'\cdot\frac{1}{r^2}\big(f(h_1+h_2) - f(h_1) + 4h_2\big)\,r\mathrm{d}r \\ &= \int_0^{+\infty}\frac{q'}{r}\big(\frac{\mathrm{d}}{\mathrm{d}r}F(h_1+h_2) - \frac{\mathrm{d}}{\mathrm{d}r}F(h_1) - h_2\cdot\frac{\mathrm{d}}{\mathrm{d}r}f(h_1) - f(h_1)\cdot\frac{\mathrm{d}}{\mathrm{d}r}h_2 + 2\frac{\mathrm{d}}{\mathrm{d}r}(h_2^2)\big)\,\mathrm{d}r \\ &= \int_0^{+\infty} r\big(\frac{q'}{r}\big)'\cdot\frac{1}{r^2}\big(F(h_1+h_2) - F(h_1) - f(h_1)\cdot h_2 + 2h_2^2\big)\,r\mathrm{d}r.\end{aligned}$$

Using (P6) in Lemma 4.6 and the elementary inequality $|F(h_1+h_2)-F(h_1)-f(h_1)h_2| \lesssim |h_1|^2|h_2|^2+|h_1|^4$ we get (4.22).

Note that as a part of this computation, we obtain

$$\big|\langle A(\lambda)h, \frac{1}{r^2}h\rangle\big| \le \frac{c_0}{\lambda}\int_0^{+\infty}\frac{1}{r^2}h^2\,r\mathrm{d}r. \tag{4.27}$$

For $r \le R\lambda$ there holds $\frac{1}{2\lambda}q''\big(\frac{r}{\lambda}\big)+\frac{1}{2r}q'\big(\frac{r}{\lambda}\big) = \frac{1}{\lambda}$ and for all r, thanks to (P4), there holds $\frac{1}{2\lambda}q''\big(\frac{r}{\lambda}\big)+\frac{1}{2r}q'\big(\frac{r}{\lambda}\big) \ge -\frac{c_0}{\lambda}$, hence

$$\langle\big(\frac{1}{2\lambda}q''\big(\frac{r}{\lambda}\big)+\frac{1}{2r}q'\big(\frac{r}{\lambda}\big)\big)h, \frac{1}{r^2}h\rangle \ge \frac{2\pi}{\lambda}\int_0^{R\lambda}\frac{1}{r^2}h^2\,r\mathrm{d}r - \frac{c_0}{\lambda}\int_0^{+\infty}\frac{1}{r^2}h^2\,r\mathrm{d}r. \tag{4.28}$$

Taking the sum of (4.27) and (4.28) we obtain

$$\big\langle A_0(\lambda)h, \frac{1}{r^2}h\big\rangle \ge -\frac{c_0}{\lambda}\int_0^{+\infty}\frac{1}{r^2}h^2\,r\mathrm{d}r + \frac{2\pi}{\lambda}\int_0^{R\lambda}\frac{1}{r^2}h^2\,r\mathrm{d}r \tag{4.29}$$

(c_0 has changed, but is still small).

Using identity (3.46) with $N=2$ we obtain, cf. the proof of (3.41),

$$\big\langle A_0(\lambda)h, \big(\partial_r^2+\frac{1}{r}\partial_r\big)h\big\rangle \le \frac{c_0}{\lambda}\int_0^{+\infty}(\partial_r h)^2\,r\mathrm{d}r - \frac{2\pi}{\lambda}\int_{r\le R\lambda}(\partial_r h)^2\,r\mathrm{d}r. \tag{4.30}$$

Taking the difference of (4.30) and (4.29) we obtain (3.41).

The proofs of (4.24) and (4.25) are similar to the proofs of (3.42) and (3.43) respectively. Instead of (3.48) we prove that $\|bA(\lambda)W_\mu\|_{L^\infty} \le \frac{c_0}{3}$, which follows from (P2) and (P3). We skip the details.

The proof of (4.26) is close to the proof of (3.44). Note that it is crucial that $f'(W_\lambda)+4$ vanishes at infinity. □

For $t \in [T, T_0]$ we define:

- the *nonlinear energy functional*

$$\begin{aligned} I(t) &:= \int \frac{1}{2}|\dot g(t)|^2 + \frac{1}{2}|\nabla g(t)|^2 - \frac{1}{r^2}\big(F(\varphi(t)+g(t)) - F(\varphi(t)) - f(\varphi(t))g(t)\big)\,\mathrm{d}x \\ &= E(\boldsymbol{\varphi}(t)+\boldsymbol{g}(t)) - E(\boldsymbol{\varphi}(t)) - \langle \mathrm{D}E(\boldsymbol{\varphi}(t), \boldsymbol{g}(t))\rangle, \end{aligned}$$

- the *localized virial functional*

$$J(t) := \int \dot g(t)\cdot A_0(\lambda(t))g(t)\,\mathrm{d}x,$$

- the *mixed energy-virial functional*

$$H(t) := I(t) + b(t)J(t).$$

4.7 Energy estimates via the mixed energy-virial functional

The remaining part of the proof is almost identical to Subsection 3.5. We will indicate the few differences.

Instead of (3.55), we obtain now

$$I'(t) \simeq -b(\lambda' - b) \cdot \langle A_0(\lambda)\Lambda W_{\underline{\lambda}}, \dot g\rangle - b \cdot \big\langle A(\lambda)g, \frac{1}{r^2}\big(f(\varphi+g) - f(\varphi) + 4g\big)\big\rangle.$$

As in the proof of Lemma 3.13, we have

$$\begin{aligned}(bJ)'(t) &\simeq b\int \dot g \cdot A_0(\lambda)(\dot g - \psi)\,r\mathrm{d}r + b\int \Big(\big(\partial_r^2 + \frac{1}{r}\partial_r\big)g \\ &+ \frac{1}{r^2}\big(f(\varphi+g) - f(\varphi)\big) - \dot\psi\Big)\cdot A_0(\lambda)g\,r\mathrm{d}r \\ &= b\int \dot g \cdot A_0(\lambda)(\dot g - \psi)\,r\mathrm{d}r + b\int \Big(\big(\partial_r^2 + \frac{1}{r}\partial_r - \frac{4}{r^2}\big)g \\ &+ \frac{1}{r^2}\big(f(\varphi+g) - f(\varphi) + 4g\big) - \dot\psi\Big)\cdot A_0(\lambda)g\,r\mathrm{d}r,\end{aligned}$$

where we recognize the terms appearing in (4.23) and (4.26). The rest of the proof applies without change. Theorem 3 follows from the argument given in Subsection 3.6.

5 Bubble-antibubble for the equivariant critical wave map equation

5.1 Notation

We use similar notation as in Section 4, with a slight modification in the definition of the norm $\mathcal{H}$:

$$\|v\|_{\mathcal{H}}^2 := 2\pi\int_0^{+\infty}\big(|\partial_r v(r)|^2 + |\frac{k}{r}v(r)|^2\big)\,r\mathrm{d}r.$$

The transformation $\widetilde v(\mathrm{e}^{i\theta}r) := \mathrm{e}^{ki\theta}v(r)$ is an isometric embedding of $\mathcal{H}$ in $\dot H^1(\mathbb{R}^2;\mathbb{R}^2)$, whose image is given by k-equivariant functions in $\dot H^1(\mathbb{R}^2;\mathbb{R}^2)$.

5.2 Linearized equation and formal computation

Linearizing $-\partial_r^2 u - \frac{1}{r}\partial_r u + \frac{k^2}{2r^2}\sin(2u)$ around $u = W$ we obtain the operator

$$L := -\partial_r^2 - \frac{1}{r}\partial_r + \frac{k^2}{r^2}\cos(4\arctan(r^k)) = -\partial_r^2 - \frac{1}{r}\partial_r + \frac{k^2}{r^2}\Big(1 - \frac{8}{(r^k + r^{-k})^2}\Big).$$

It has a one-dimensional kernel spanned by ΛW. We fix $\mathcal{Z} \in C_0^\infty((0,+\infty))$ such that

$$\int_0^{+\infty}\mathcal{Z}(r)\cdot\Lambda W(r)\frac{\mathrm{d}r}{r} > 0.$$

Lemma 5.1. *For all $V(r) \in C^\infty((0,+\infty))$ such that $\int_0^{+\infty}\Lambda W(r)\cdot V(r)\,r\mathrm{d}r = 0$, $|V(r)| \lesssim r^k$ for small r and $|V(r)| \lesssim r^{-k}$ for large r, then there exists a function $U(r) \in C^\infty((0,+\infty))$ such that*

$$LU = V, \tag{5.1}$$

$$|U(r)| \lesssim r^k, \quad |\partial_r U(r)| \lesssim r^{k-1}, \quad |\partial_r^2 U(r)| \lesssim r^{k-2} \qquad \text{for } r \text{ small}, \tag{5.2}$$

$$|U(r)| \lesssim r^{-k}, \quad |\partial_r U(r)| \lesssim r^{-k-1}, \quad |\partial_r^2 U(r)| \lesssim r^{-k-2} \qquad \textit{for } r \textit{ large}, \tag{5.3}$$

$$\int \mathcal{Z}(r) \cdot U(r) \frac{\mathrm{d}r}{r} = 0. \tag{5.4}$$

Proof. It is easy to check that the operator L factorizes as follows:

$$L = -\partial_r^2 - \frac{1}{r}\partial_r + \frac{k^2}{r^2}\Big(1 - \frac{8}{(r^k + r^{-k})^2}\Big) = \Big(-\partial_r - \frac{1}{r} - \frac{\Lambda W'(r)}{\Lambda W(r)}\Big)\Big(\partial_r - \frac{\Lambda W'(r)}{\Lambda W(r)}\Big), \tag{5.5}$$

hence we can invert it explicitely using twice the variation of constants formula. Define $U_1 \in C^\infty((0, +\infty))$ by

$$U_1(r) := \frac{1}{r\Lambda W(r)} \int_0^r V(\rho)\Lambda W(\rho)\rho \,\mathrm{d}\rho.$$

It solves the equation $\Big(-\partial_r - \frac{1}{r} - \frac{\Lambda W'(r)}{\Lambda W(r)}\Big)U_1(r) = V(r)$. Since $|V(r)| \lesssim r^k$ and $\Lambda W(r) \sim r^k$ for small r, we have

$$|U_1(r)| \lesssim r^{-1-k+k+k+1+1} = r^{k+1}, \qquad \text{small } r. \tag{5.6}$$

From the crucial assumption $\int_0^{+\infty} V(\rho)\Lambda W(\rho)\rho\,\mathrm{d}\rho = 0$ we get

$$|U_1(r)| = \Big|\frac{1}{r\Lambda W(r)} \int_r^{+\infty} V(\rho)\Lambda W(\rho)\rho\,\mathrm{d}\rho\Big| \lesssim r^{-k+1}, \qquad \text{large } r. \tag{5.7}$$

From the differential equation we get also $|\partial_r U_1(r)| \lesssim r^k$ for small r and $|\partial_r U_1(r)| \lesssim r^{-k}$ for large r. Now we define $U \in C^\infty((0, +\infty))$ by the formula

$$U(r) := \Lambda W(r) \int_0^r \frac{U_1(\rho)}{\Lambda W(\rho)}\,\mathrm{d}\rho.$$

It solves $\Big(\partial_r - \frac{\Lambda W'(r)}{\Lambda W(r)}\Big)U(r) = U_1(r)$, hence (5.5) yields (5.1). Using (5.6) and (5.7), one can check that $|U(r)| \lesssim r^{k+2}$ for small r and $|U(r)| \lesssim r^{-k+2}$ for large r. The differential equations yield $|\partial_r U(r)| \lesssim r^{k+1}$ and $\partial_r^2 U(r)| \lesssim r^k$ for small r, as well as $|\partial_r U(r)| \lesssim r^{-k+1}$ and $|\partial_r^2 U(r)| \lesssim r^{-k}$ for large r. Adding to U a suitable multiple of ΛW we obtain (5.4). Since $|\Lambda W(r)| \lesssim r^k$, $|\partial_r \Lambda W(r)| \lesssim r^{k-1}$, $|\partial_r^2 \Lambda W(r)| \lesssim r^{k-2}$ for small r and $|\Lambda W(r)| \lesssim r^{-k}$, $|\partial_r \Lambda W(r)| \lesssim r^{-k-1}$, $|\partial_r^2 \Lambda W(r)| \lesssim r^{-k-2}$ for large r, (5.2) and (5.3) still hold. ☐

We study solutions behaving like $\boldsymbol{u}(t) \simeq -\boldsymbol{W} + \boldsymbol{W}_{\lambda(t)}$ with $\lambda(t) \to 0$ as $t \to -\infty$. We expand

$$\boldsymbol{u}(t) = -\boldsymbol{W} + \boldsymbol{U}^{(0)}_{\lambda(t)} + b(t) \cdot \boldsymbol{U}^{(1)}_{\lambda(t)} + b(t)^2 \cdot \boldsymbol{U}^{(2)}_{\lambda(t)},$$

with $b(t) = \lambda'(t)$, $\boldsymbol{U}^{(0)} := (W, 0)$ and $\boldsymbol{U}^{(1)} := (0, -\Lambda W)$. As in Subsection 4.2, in the region $r \leq \sqrt{\lambda}$ we arrive at

$$\partial_r^2 u + \frac{1}{r}\partial_r u - \frac{k^2}{2r^2}\sin(2u) = -\frac{b^2}{\lambda}(LU^{(2)})_{\underline{\lambda}} - \frac{1}{r^2} \cdot \frac{8k^2}{((r/\lambda)^k + (r/\lambda)^{-k})^2} W + \mathrm{lot}.$$

Using the fact that $W(r) \sim 2r^k$ for small r we obtain

$$\partial_r^2 u + \frac{1}{r}\partial_r u - \frac{k^2}{2r^2}\sin(2u) = -\frac{b^2}{\lambda}(LU^{(2)})_{\underline{\lambda}} - \frac{16k^2 r^{k-2}}{((r/\lambda)^k + (r/\lambda)^{-k})^2} + \mathrm{lot},$$

thus, after rescaling,

$$LU^{(2)} = -\Lambda_0\Lambda W + \frac{\lambda}{b^2}\big(b'\Lambda W - \lambda^{k-1}\cdot\frac{16k^2r^{k-2}}{(r^k+r^{-k})^2}\big).$$

It is not difficult to check (using for example the residue theorem) that

$$\int_0^{+\infty}\Lambda W(r)\cdot\frac{16k^2r^{k-2}}{(r^k+r^{-k})^2}\,r\mathrm{d}r = \frac{4k^2}{\pi}\sin\big(\frac{\pi}{k}\big)\cdot\int\Lambda W(r)^2\,r\mathrm{d}r,$$

hence the correct choice (that is, such that Lemma 5.1 allows to invert L) of the formal parameter equations is

$$\lambda'(t) = b(t), \qquad b'(t) = \frac{4k^2}{\pi}\sin\big(\frac{\pi}{k}\big)\lambda(t)^{k-1} = \frac{k}{2}\Big(\frac{2\kappa}{k-2}\Big)^k\lambda(t)^{k-1},$$

where $\kappa := \frac{k-2}{2}\cdot\big(\frac{8k}{\pi}\sin\big(\frac{\pi}{k}\big)\big)^{\frac{1}{k}}$. This system has a solution

$$(\lambda_{\mathrm{app}}(t), b_{\mathrm{app}}(t)) = \big(\frac{k-2}{2\kappa}(\kappa|t|)^{-\frac{2}{k-2}}, (\kappa|t|)^{-\frac{k}{k-2}}\big), \qquad t\leq T_0<0.$$

5.3 Bounds on the error of the ansatz

Let $I = [T, T_0]$ be the time interval, $T\leq T_0<0$, $|T_0|$ large. Let $\lambda(t)$ and $\mu(t)$ be C^1 functions on $[T, T_0]$ such that

$$\begin{aligned}\lambda(T) &= \frac{k-2}{\kappa}(\kappa|T|)^{-\frac{2}{k-2}}, \quad \mu(T) = 1,\\ \frac{k-2}{2\kappa}(\kappa|t|)^{-\frac{2}{k-2}} \leq \lambda(t) &\leq \frac{2(k-2)}{\kappa}(\kappa|t|)^{-\frac{2}{k-2}}, \qquad \frac{1}{2}\leq\mu(t)\leq 2.\end{aligned} \tag{5.8}$$

Let $P(r)$ and $Q(r)$ by the functions obtained in Lemma 5.1 for $V(r) = \frac{k}{2}\big(\frac{2\kappa}{k-2}\big)^k\Lambda W(r) - \frac{16k^2r^{k-2}}{(r^k+r^{-k})^2}$ and $V(r) = -\Lambda_0\Lambda W(r)$ respectively. We define the approximate solution by the formula

$$\begin{aligned}\varphi(t) &:= -W_{\mu(t)} + W_{\lambda(t)} + S(t),\\ \dot\varphi(t) &:= -b(t)\Lambda W_{\underline{\lambda(t)}}.\end{aligned}$$

where

$$\begin{aligned}b(t) &:= (\kappa|T|)^{-\frac{k}{k-2}} + \Big(\frac{2\kappa}{k-2}\Big)^k\int_T^t\frac{\lambda(\tau)^{k-1}}{\mu(\tau)^k}\,\mathrm{d}\tau, && \text{for } t\in[T,T_0], \qquad (5.9)\\ S(t) &:= \frac{\lambda(t)^k}{\mu(t)^k}P_{\lambda(t)} + b(t)^2Q_{\lambda(t)}, && \text{for } t\in[T,T_0].\end{aligned}$$

The definition of $b(t)$ and (5.8) yield

$$b(t)\sim|t|^{-\frac{k}{k-2}},$$

with a constant depending only on k.

Since $P, Q\in\mathcal{H}$, there holds

$$\|S(t)\|_{\mathcal{H}}\lesssim|t|^{-\frac{2k}{k-2}}. \tag{5.10}$$

Note also that

$$\int \mathcal{Z}_{\underline{\lambda}} \cdot S(t)\frac{\mathrm{d}r}{r} = 0.$$

We denote $f(u) := -\frac{k^2}{2}\sin(2u)$ (hence $f'(u) = -k^2\cos(2u)$) and

$$\begin{aligned}\boldsymbol{\psi}(t) = (\psi(t), \dot\psi(t)) &:= \partial_t\boldsymbol{\varphi}(t) - \mathrm{D}E(\boldsymbol{\varphi}(t)) \\ &= \big(\partial_t\varphi(t) - \dot\varphi(t), \partial_t\dot\varphi(t) - (\partial_r^2\varphi(t) + \frac{1}{r}\partial_r\varphi(t) + \frac{1}{r^2}f(\varphi(t)))\big).\end{aligned}$$

We point out that usually, in the context of equivariant wave maps, $f(u)$ has the opposite sign. We chose the sign which is more coherent with the traditional notation for (1.1).

Lemma 5.2. *Suppose that for $t \in [T, T_0]$ there holds $|\lambda'(t)| \lesssim |t|^{-\frac{k}{k-2}}$ and $|\mu'(t)| \lesssim |t|^{-\frac{k}{k-2}}$. Then*

$$\|\psi(t) - \mu'(t)\frac{1}{\mu(t)}\Lambda W_{\mu(t)} + (\lambda'(t) - b(t))\frac{1}{\lambda(t)}\Lambda W_{\lambda(t)}\|_{\mathcal{H}} \lesssim |t|^{-\frac{2k-1}{k-2}},$$

$$\|\dot\psi(t) - \frac{b(t)}{\lambda(t)}(\lambda'(t) - b(t))\Lambda_0\Lambda W_{\underline{\lambda(t)}}\|_{L^2} \lesssim |t|^{-\frac{2k-1}{k-2}}, \tag{5.11}$$

$$\|(-\partial_r^2 - \frac{1}{r}\partial_r - \frac{1}{r^2}f'(\varphi(t)))\psi(t)\|_{\mathcal{H}^*} \lesssim |t|^{-\frac{2k-1}{k-2}}. \tag{5.12}$$

Proof. From the definition of $\boldsymbol{\psi}$ we get

$$\psi + \mu'\Lambda W_{\underline{\mu}} + (\lambda' - b)\Lambda W_{\underline{\lambda}} = -k\frac{\lambda^k}{\mu^{k+1}}\mu' P_\lambda + k\frac{\lambda^k}{\mu^k}\lambda' P_{\underline{\lambda}} - \frac{\lambda^k}{\mu^k}\lambda'\Lambda P_{\underline{\lambda}} + 2\lambda b b' Q_{\underline{\lambda}} - b^2\lambda'\Lambda Q_{\underline{\lambda}}.$$

The first term has size $\lesssim |t|^{-\frac{3k}{k-2}} \ll |t|^{-\frac{2k-1}{k-2}}$ in $\mathcal{H}$. The other terms have size $\lesssim |t|^{-\frac{3k-2}{k-2}} \ll |t|^{-\frac{2k-1}{k-2}}$ in $\mathcal{H}$.

In order to prove (5.11), we treat separately the regions $r \le \sqrt\lambda$ and $r \ge \sqrt\lambda$. First we will show that

$$\Big\|\frac{1}{r^2}\big(f(\varphi) - f(W_\lambda) + 2\frac{r^k}{\mu^k}f'(W_\lambda) - \frac{\lambda^k}{\mu^k}f'(W_\lambda)P_\lambda - b^2 f'(W_\lambda)Q_\lambda\big)\Big\|_{L^2(r\le\sqrt\lambda)} \lesssim |t|^{-\frac{2k-1}{k-2}}. \tag{5.13}$$

We have an elementary pointwise inequality

$$|f(\varphi) - f(W_\lambda) - f'(W_\lambda)(-W_\mu + S)| \lesssim |-W_\mu + S|^2, \tag{5.14}$$

with a constant depending only on k. We have $|W_\mu| \lesssim r^k$ and $|S| \lesssim (b^2 + \lambda^k)\cdot\big(\frac{r}{\lambda}\big)^k \lesssim r^k$, hence

$$\Big\|\frac{1}{r^2}|-W_\mu + S|^2\Big\|_{L^2(r\le\sqrt\lambda)} \lesssim \Big(\int_0^{\sqrt\lambda}\big(\frac{1}{r^2}\cdot r^{2k}\big)^2 r\,\mathrm{d}r\Big)^{\frac12} \sim \lambda^{\frac{2k-1}{2}} \lesssim |t|^{-\frac{2k-1}{k-2}},$$

which means that the right hand side in (5.14) is negligible. From the well-known fact that $|\arctan(z) - z| \lesssim |z|^3$ for small z we get $|W_\mu(r) - 2(r/\mu)^k| \lesssim r^{3k}$. This implies

$$\Big\|\frac{1}{r^2}f'(W_\lambda)\big(W_\mu - 2\frac{r^k}{\mu^k}\big)\Big\|_{L^2} \lesssim \Big(\int_0^{\sqrt\lambda} r^{6k-4}\,r\,\mathrm{d}r\Big)^{\frac12} \lesssim \lambda^{\frac{3k-1}{2}} \lesssim |t|^{-\frac{3k-1}{k-2}} \ll |t|^{-\frac{2k-1}{k-2}}.$$

This proves (5.13) (see the definition of S).

We have $\left|W_\mu'(r) - \frac{2kr^{k-1}}{\mu^k}\right| \lesssim r^{3k-1}$ and $\left|W_\mu''(r) - \frac{2(k^2-k)r^{k-2}}{\mu^k}\right| \lesssim r^{3k-2}$ for small r, which implies $\left|\left(\partial_r^2 + \frac{1}{r}\partial_r\right)W_\mu(r) - \frac{2k^2r^{k-2}}{\mu^k}\right| \lesssim r^{3k-2}$, hence

$$\left\|\left(\partial_r^2 + \frac{1}{r}\partial_r\right)W_\mu - \frac{2k^2r^{k-2}}{\mu^k}\right\|_{L^2(r\leq\sqrt{\lambda})} \lesssim \left(\int_0^{\sqrt{\lambda}} r^{6k-4}\,r\mathrm{d}r\right)^{\frac{1}{2}} \ll |t|^{-\frac{2k-1}{k-2}}. \tag{5.15}$$

Since $W_\lambda''(r) + \frac{1}{r}W_\lambda'(r) + \frac{1}{r^2}f(W_\lambda(r)) = 0$, from (5.13), (5.15) and the definition of $\varphi(t)$ we have

$$\begin{aligned}\Big\|\Big(\big(\partial_r^2 + \frac{1}{r}\partial_r\big)\varphi + \frac{1}{r^2}f(\varphi)\Big) - \Big(-\frac{2k^2r^{k-2}}{\mu^k} - \frac{2r^{k-2}}{\mu^k}f'(W_\lambda) - \frac{\lambda^{k-1}}{\mu^k}(LP)_{\underline{\lambda}}\\ - \frac{b^2}{\lambda}(LQ)_{\underline{\lambda}}\Big)\Big\|_{L^2(r\leq\sqrt{\lambda})} \lesssim \mathrm{e}^{-\frac{3}{2}\kappa|t|}.\end{aligned} \tag{5.16}$$

By the definition of P there holds $LP = -\frac{16k^2r^{k-2}}{(r^k+r^{-k})^2} + \frac{k}{2}\left(\frac{2\kappa}{k-2}\right)^k\Lambda W = -2k^2r^{k-2} - 2r^{k-2}f'(W) + \frac{k}{2}\left(\frac{2\kappa}{k-2}\right)^k\Lambda W$ and by the definition of Q there holds $LQ = -\Lambda_0\Lambda W$, hence we can rewrite (5.16) as

$$\Big\|\Big(\big(\partial_r^2 + \frac{1}{r}\partial_r\big)\varphi + \frac{1}{r^2}f(\varphi)\Big) - \Big(-\frac{k}{2}\big(\frac{2\kappa}{k-2}\big)^k\frac{\lambda^{k-1}}{\mu^k}\Lambda W_{\underline{\lambda}} + \frac{b^2}{\lambda}\Lambda_0\Lambda W_{\underline{\lambda}}\Big)\Big\|_{L^2(r\leq\sqrt{\lambda})} \lesssim \mathrm{e}^{-\frac{3}{2}\kappa|t|},$$

which is precisely (5.11) restricted to $r \leq \sqrt{\lambda}$, cf. (2.25).

Consider the region $r \geq \sqrt{\lambda}$. We have $\varphi = (\pi - W_\mu) + (W_\lambda - \pi) + S$, hence elementary pointwise inequalities yield

$$|f(\varphi) - f(\pi - W_\mu) - f'(\pi - W_\mu)(W_\lambda - \pi + S)| \lesssim |W_\lambda - \pi + S|^2.$$

From this and the relations $f(\pi - W_\mu) = -f(W_\mu)$, $f'(\pi - W_\mu) = f'(W_\mu)$, we obtain a pointwise bound

$$|f(\varphi) + f(W_\mu) + f'(W_\mu)(\pi - W_\lambda)| \lesssim |S| + |\pi - W_\lambda|^2. \tag{5.17}$$

Since $|S(r)| \lesssim b^2 + \frac{\lambda^k}{\mu^k} \sim |t|^{-\frac{2k}{k-2}}$, we have

$$\left\|\frac{1}{r^2}S\right\|_{L^2(r\geq\sqrt{\lambda})} \lesssim |t|^{-\frac{2k}{k-2}}\left(\int_{\sqrt{\lambda}}^{+\infty} r^{-4}\,r\mathrm{d}r\right)^{\frac{1}{2}} \lesssim \frac{|t|^{-\frac{2k}{k-2}}}{\sqrt{\lambda}} \lesssim |t|^{-\frac{2k-1}{k-2}}. \tag{5.18}$$

There holds $|\pi - W_\lambda(r)|^2 = |\pi - 2\arctan((r/\lambda)^k)|^2 = |2\arctan((\lambda/r)^k)|^2 \lesssim \frac{\lambda^{2k}}{r^{2k}}$, hence

$$\left\|\frac{1}{r^2}|\pi - W_\lambda(r)|^2\right\|_{L^2(r\geq\sqrt{\lambda})} \lesssim \lambda^{2k}\left(\int_{\sqrt{\lambda}}^{+\infty} r^{-4k-4}\,r\mathrm{d}r\right)^{\frac{1}{2}} \sim \lambda^{2k}\lambda^{-\frac{2k+1}{2}} \sim |t|^{-\frac{2k-1}{k-2}}. \tag{5.19}$$

Recall that $f'(W_\mu) = -k^2(1 - 8((r/\mu)^k + (r/\mu)^{-k})^{-2})$, hence $|f'(W_\mu) + k^2| \lesssim r^k$. We also have (by a standard asymptotic expansion of arctan) $|\pi - W_\lambda| \lesssim \frac{\lambda^k}{r^k}$ and $|\pi - W_\lambda - 2\frac{\lambda^k}{r^k}| \lesssim \frac{\lambda^{3k}}{r^{3k}}$, hence

$$\left\|\frac{1}{r^2}f'(W_\mu)(\pi - W_\lambda) + \frac{2k^2\lambda^k}{r^{k+2}}\right\|_{L^2(r\geq\sqrt{\lambda})} \lesssim \left\|\frac{\lambda^k}{r^2} + \frac{\lambda^{3k}}{r^{3k+2}}\right\|_{L^2(r\geq\sqrt{\lambda})} \lesssim |t|^{-\frac{2k-1}{k-2}}. \tag{5.20}$$

Inserting (5.18), (5.19) and (5.20) into (5.17) we obtain

$$\left\|\frac{1}{r^2}f(\varphi) + \frac{1}{r^2}f(W_\mu) - \frac{2k^2\lambda^k}{r^{k+2}}\right\|_{L^2(r\geq\sqrt{\lambda})} \lesssim |t|^{-\frac{2k-1}{k-2}}. \tag{5.21}$$

A direct computation shows that for $r \geq \sqrt{\lambda}$ there holds $(\partial_r^2+\frac{1}{r}\partial_r)W_\lambda(r) = -\frac{2k^2\lambda^k}{r^{k+2}}+O\big(\frac{\lambda^{3k}}{r^{3k+2}}\big)$, hence

$$\|(\partial_r^2 + \frac{1}{r}\partial_r)W_\lambda + \frac{2k^2\lambda^k}{r^{k+2}}\|_{L^2(r\geq\sqrt{\lambda})} \lesssim \lambda^{\frac{3k-1}{2}} \ll |t|^{-\frac{2k-1}{k-2}}. \tag{5.22}$$

The same computation as in (4.15) yields $\|(\partial_r^2 + \frac{1}{r}\partial_r)S\|_{L^2(r\geq\sqrt{\lambda})} \lesssim \frac{b^2}{\sqrt{\lambda}} + \frac{\lambda^k}{\sqrt{\lambda}} \lesssim |t|^{-\frac{2k-1}{k-2}}$. Together with (5.21) and (5.22) this proves that

$$\|(\partial_r^2 + \frac{1}{r}\partial_r)\varphi + \frac{1}{r^2}f(\varphi)\|_{L^2(r\geq\sqrt{\lambda})} \lesssim |t|^{-\frac{2k-1}{k-2}}.$$

Since $\|\Lambda W_{\underline{\lambda}}\|_{L^2(r\geq\sqrt{\lambda})} + \|\Lambda_0\Lambda W_{\underline{\lambda}}\|_{L^2(r\geq\sqrt{\lambda})} \lesssim \big(\int_{1/\sqrt{\lambda}}^{+\infty}\frac{1}{r^{2k}}\,r\mathrm{d}r\big)^{\frac{1}{2}} \lesssim \lambda^{\frac{k-1}{2}}$, the other terms appearing in (5.11) are $\lesssim |t|^{-\frac{3k-3}{k-2}} \ll |t|^{-\frac{2k-1}{k-2}}$. This finishes the proof of (5.11).

The proof of (5.12) is very similar to the proof of (2.19), so we will not give all the details. We have $\big|\frac{\lambda'-b}{\lambda}\big| \lesssim |t|^{-1}$ and $\big|\frac{\mu'}{\mu}\big| \lesssim |t|^{-1}$, hence it suffices to check that $\big\|\frac{1}{r^2}(f'(\varphi) - f'(W_\lambda))\Lambda W_\lambda\big\|_{\mathcal{H}^*} \lesssim |t|^{-\frac{k+1}{k-2}}$ and $\big\|\frac{1}{r^2}(f'(\varphi) - f'(W_\mu))\Lambda W_\mu\big\|_{\mathcal{H}^*} \lesssim |t|^{-\frac{k+1}{k-2}}$, which boils down to $\big\|\frac{1}{r^2}W_\mu\cdot\Lambda W_\lambda\big\|_{L^1} \lesssim |t|^{-\frac{k+1}{k-2}}$ and $\big\|\frac{1}{r^2}(\pi - W_\lambda)\cdot\Lambda W_\mu\big\|_{L^1} \lesssim |t|^{-\frac{k+1}{k-2}}$, see the proof of (4.6). In both cases we treat separately $r \leq 1$ and $r \geq 1$:

$$\begin{aligned}
\big\|\frac{1}{r^2}W_\mu\cdot\Lambda W_\lambda\big\|_{L^1(r\leq 1)}\|r^{k-2}\Lambda W_\lambda\|_{L^1(r\leq 1)} = \lambda^k\|\Lambda W\|_{L^1(r\leq 1/\lambda)} \lesssim \lambda^k|\log\lambda| \ll |t|^{-\frac{k+1}{k-2}},\\
\big\|\frac{1}{r^2}W_\mu\cdot\Lambda W_\lambda\big\|_{L^1(r\geq 1)}\|\Lambda W\|_{L^\infty(r\geq 1)} \lesssim \lambda^k \ll |t|^{-\frac{k+1}{k-2}},\\
\big\|\frac{1}{r^2}(\pi - W_\lambda)\cdot\Lambda W_\mu\big\|_{L^1(r\leq 1)} \lesssim \big\|r^{k-2}\arctan\big((\lambda/r)^k\big)\big\|_{L^1(r\leq 1)}\\
= \lambda^k\|r^{k-2}\arctan(r^{-k})\|_{L^1(r\leq 1/\lambda)} \lesssim \lambda^k|\log\lambda| \ll |t|^{-\frac{k+1}{k-2}},\\
\big\|\frac{1}{r^2}(\pi - W_\lambda)\cdot\Lambda W_\mu\big\|_{L^1(r\geq 1)} \lesssim \|\pi - W_\lambda\|_{L^\infty(r\geq 1)} \lesssim \lambda^k \ll |t|^{-\frac{k+1}{k-2}}.
\end{aligned}$$

□

5.4 Modulation

As in Subsection 3.2, we define $\boldsymbol{g}(t) := \boldsymbol{u}(t) - \boldsymbol{\varphi}(t)$ with modulation parameters $\lambda(t)$ and $\mu(t)$ which satisfy

$$\begin{gathered}
\big\langle\frac{1}{\lambda(t)}\mathcal{Z}_{\underline{\lambda(t)}}, g(t)\big\rangle = 0, \quad \big|\big\langle\frac{1}{\mu(t)}\mathcal{Z}_{\underline{\mu(t)}}, g(t)\big\rangle\big| \lesssim c|t|^{-\frac{k+1}{k-2}},\\
|\lambda'(t) - b(t)| + |\mu'(t)| \lesssim \|\boldsymbol{g}(t)\|_{\mathcal{E}} + c|t|^{-\frac{k+1}{k-2}},
\end{gathered} \tag{5.23}$$

with a constant c arbitrarily small. The initial data are

$$\boldsymbol{u}(T) = \big(-W + W_{\frac{k-2}{2\kappa}(\kappa|t|)^{-\frac{2}{k-2}}}, -(\kappa|t|)^{-\frac{k}{k-2}}\Lambda W_{\underline{\frac{k-2}{2\kappa}(\kappa|t|)^{-\frac{2}{k-2}}}}\big). \tag{5.24}$$

The equivalent of Proposition 3.3 can by formulated as follows.

Proposition 5.3. *There exist constants $C_0 > 0$ and $T_0 < 0$ with the following property Let $T < T_1 < T_0$ and suppose that $\boldsymbol{u}(t) = \boldsymbol{\varphi}(t) + \boldsymbol{g}(t) \in C([T, T_1]; X^1 \times X^0)$ is a solution of* (1.3) *with initial data* (5.24) *such that for $t \in [T, T_1]$ condition* (5.8) *is satisfied and*

$$\|\boldsymbol{g}(t)\|_{\mathcal{E}} \leq C_0|t|^{-\frac{k+1}{k-2}}. \tag{5.25}$$

Then for $t \in [T, T_1]$ there holds

$$\|\boldsymbol{g}(t)\|_{\mathcal{E}} \leq \frac{1}{2}C_0|t|^{-\frac{k+1}{k-2}}, \tag{5.26}$$

$$\big|\lambda(t) - \frac{k-2}{2\kappa}(\kappa|t|)^{-\frac{2}{k-2}}\big| + |\mu(t) - 1| \lesssim C_0|t|^{-\frac{3}{k-2}}. \tag{5.27}$$

The rest of this section is devoted to the proof of this bootstrap estimate. Proposition 5.3 allows to construct a uniformly controlled sequence of solutions close to the ansatz and pass to a weak limit, thereby proving Theorem 2, see Subsection 3.6.

5.5 Coercivity

Recall that $f'(W) = -k^2\Big(1 - \frac{8}{(r^k + r^{-k})^2}\Big)$.

Lemma 5.4. *There exist constants $c, C > 0$ such that*

- *for all $g \in \mathcal{H}$ there holds*

$$\begin{aligned}&\int_0^{+\infty}\big((g')^2 + \frac{k^2}{r^2}g^2\big)\,r\mathrm{d}r - \int_0^{+\infty}\frac{k^2}{r^2}\cdot\frac{8}{(r^k + r^{-k})^2}g^2\,r\mathrm{d}r \\ &\geq c\int_0^{+\infty}\big((g')^2 + \frac{k^2}{r^2}g^2\big)\,r\mathrm{d}r - C\Big(\int_0^{+\infty}\mathcal{Z}\cdot g\frac{\mathrm{d}r}{r}\Big)^2,\end{aligned}$$

- *if $r_1 > 0$ is large enough, then for all $g \in \mathcal{H}$ there holds*

$$\begin{aligned}&(1-2c)\int_0^{r_1}\big((g')^2 + \frac{k^2}{r^2}g^2\big)\,r\mathrm{d}r + c\int_{r_1}^{+\infty}\big((g')^2 + \frac{k^2}{r^2}g^2\big)\,r\mathrm{d}r \\ &- \int_0^{+\infty}\frac{k^2}{r^2}\cdot\frac{8}{(r^k + r^{-k})^2}g^2\,r\mathrm{d}r \geq -C\Big(\int_0^{+\infty}\mathcal{Z}\cdot g\frac{\mathrm{d}r}{r}\Big)^2,\end{aligned} \tag{5.28}$$

- *if $r_2 > 0$ is small enough, then for all $g \in \mathcal{H}$ there holds*

$$\begin{aligned}&(1-2c)\int_{r_2}^{+\infty}\big((g')^2 + \frac{k^2}{r^2}g^2\big)\,r\mathrm{d}r + c\int_0^{r_2}\big((g')^2 + \frac{k^2}{r^2}g^2\big)\,r\mathrm{d}r \\ &- \int_0^{+\infty}\frac{k^2}{r^2}\cdot\frac{8}{(r^k + r^{-k})^2}g^2\,r\mathrm{d}r \geq -C\Big(\int_0^{+\infty}\mathcal{Z}\cdot g\frac{\mathrm{d}r}{r}\Big)^2,\end{aligned}$$

Proof. A change of variable $\widetilde{g}(x) := g(\mathrm{e}^{x/k})$, $\widetilde{\mathcal{Z}}(x) := \mathcal{Z}(\mathrm{e}^{x/k})$ reduces the problem to the study of the quadratic form associated with the classical operator $-\frac{\mathrm{d}^2}{\mathrm{d}x^2} + (1 - 2\,\mathrm{sech}^2)$ and it suffices to repeat the proof of Lemma 4.4. □

Lemma 4.5 applies verbatim to the case under consideration.

5.6 Bootstrap

We use the operators $A(\lambda)$ and $A_0(\lambda)$ from Subsection 4.6. We define $I(t)$, $J(t)$ and $H(t)$ by the same formulas as in Subsection 4.6.

Lemma 5.5. *The operators $A(\lambda)$ and $A_0(\lambda)$ have the following properties:*

- *for all $\lambda > 0$ and $h_1, h_2 \in X^1$ there holds*

$$\begin{aligned}&\big|\big\langle A(\lambda)h_1, \frac{1}{r^2}\big(f(h_1+h_2) - f(h_1) - f'(h_1)h_2\big)\big\rangle \\ &+ \big\langle A(\lambda)h_2, \frac{1}{r^2}\big(f(h_1+h_2) - f(h_1) + k^2 h_2\big)\big\rangle\big| \leq \frac{c_0}{\lambda}\cdot\|h_2\|_{\mathcal{H}}^2,\end{aligned} \tag{5.29}$$

with a constant c_0 arbitrarily small,

- *for all $h \in X^1$ there holds*

$$\big\langle A_0(\lambda)h, \big(\partial_r^2 + \frac{1}{r}\partial_r - \frac{k^2}{r^2}\big)h\big\rangle \leq \frac{c_0}{\lambda}\|h\|_{\mathcal{H}}^2 - \frac{2\pi}{\lambda}\int_0^{R\lambda}\big((\partial_r h)^2 + \frac{k^2}{r^2}h^2\big)\,r\mathrm{d}r, \tag{5.30}$$

- *assuming (2.7), for any $c_0 > 0$ there holds*

$$\|\Lambda_0\Lambda W_{\underline{\lambda(t)}} - A_0(\lambda(t))\Lambda W_{\lambda(t)}\|_{L^2} \leq c_0, \tag{5.31}$$

$$\|\dot\varphi(t) + b(t)\cdot A(\lambda(t))\varphi(t)\|_{L^\infty} \leq c_0|t|^{-1}, \tag{5.32}$$

$$\begin{aligned}\Big|\int_0^{+\infty}&\frac{1}{2}\big(q''\big(\frac{r}{\lambda}\big) + \frac{\lambda}{r}q'\big(\frac{r}{\lambda}\big)\big)\frac{1}{r^2}\big(f(\varphi+g) - f(\varphi) + k^2 g\big)g\,r\mathrm{d}r \\ &- \int_0^{+\infty}\frac{1}{r^2}\big(f'(W_\lambda) + k^2\big)g^2\,r\mathrm{d}r\Big| \leq c_0 C_0^2|t|^{-\frac{2k+2}{k-2}},\end{aligned} \tag{5.33}$$

provided that the constant R in the definition of $q(r)$ is chosen large enough.

□

The proof is almost identical to the proof of Lemma 4.7 and we will skip it.

Lemma 5.6. *Let $c_1 > 0$. If C_0 is sufficiently large, then there exists a function $q(x)$ and $T_0 < 0$ with the following property. If $T_1 < T_0$ and (5.8), (5.25) hold for $t \in [T, T_1]$, then for $t \in [T, T_1]$ there holds*

$$H'(t) \leq c_1\cdot C_0^2|t|^{-\frac{3k}{k-2}}. \tag{5.34}$$

Proof. We follow the lines of the proof of Lemma 3.13. We have

$$I'(t) = \big\langle\big(\partial_r^2 + \frac{1}{r}\partial_r + \frac{1}{r^2}f'(\varphi)\big)\psi, g\big\rangle - \langle\dot\psi, g\rangle - \big\langle\dot\varphi, \frac{1}{r^2}(f(\varphi+g) - f(\varphi) - f'(\varphi)g)\big\rangle.$$

The first term is $\lesssim C_0|t|^{-\frac{3k}{k-2}}$, hence negligible (by enlarging C_0 if necessary). Inequality (5.11) implies that the second term can be replaced by $-\frac{b}{\lambda}(\lambda' - b)\langle\Lambda_0\Lambda W_{\underline{\lambda}}, \dot g\rangle$, which in turn can be replaced by $-b(\lambda' - b)\langle A_0(\lambda)\Lambda W_{\underline{\lambda}}, \dot g\rangle$, thanks to (5.31). From (5.32) we infer that the third term can be replaced by $b\cdot\big\langle A(\lambda)\varphi, \frac{1}{r^2}(f(\varphi+g) - f(\varphi) - f'(\varphi)g)\big\rangle$ (indeed, $\big\|\frac{1}{r^2}(f(\varphi+g) - f(\varphi) - f'(\varphi)g)\big\|_{L^1} \lesssim \int_0^{+\infty}\frac{1}{r^2}g^2\,r\mathrm{d}r \lesssim \|g\|_{\mathcal{H}}^2$). Using formula (5.29) with $h_1 = \varphi$ and $h_2 = g$ we obtain

$$I'(t) \simeq -b(\lambda' - b)\cdot\langle A_0(\lambda)\Lambda W_{\underline{\lambda}}, \dot g\rangle - b\cdot\big\langle A(\lambda)g, \frac{1}{r^2}(f(\varphi+g) - f(\varphi) + k^2 g)\big\rangle.$$

As in the proof of Lemma 3.13, we obtain

$$\begin{aligned}(bJ)'(t) &\simeq b(\lambda' - b)\cdot\langle A_0(\lambda)\Lambda W_{\underline{\lambda}}, \dot g\rangle \\ &+ b\int\big(\big(\partial_r^2 + \frac{1}{r}\partial_r - \frac{k^2}{r^2}\big)g + \frac{1}{r^2}\big(f(\varphi+g) - f(\varphi) + k^2 g\big)\big)\cdot A_0(\lambda)g\,r\mathrm{d}r,\end{aligned}$$

hence

$$H'(t) \simeq b\int \big((\partial_r^2 + \frac{1}{r}\partial_r - \frac{k^2}{r^2})g\big)\cdot A_0(\lambda)g\,r\mathrm{d}r$$
$$- \frac{b}{\lambda}\int_0^{+\infty}\frac{1}{2}\big(q''(\frac{r}{\lambda}) + \frac{\lambda}{r}q'(\frac{r}{\lambda})\big)\frac{1}{r^2}\big(f(\varphi+g) - f(\varphi) + k^2 g\big)g\,r\mathrm{d}r.$$

and the conclusion follows from (5.30), (5.33) and (5.28). □

Proof of Proposition 5.3. We first show (5.27). From (5.23) and (5.25) we obtain

$$|\mu(t) - 1| = |\mu(t) - \mu(T)| \lesssim \int_{-\infty}^t C_0|t|^{-\frac{k+1}{k-2}}\,\mathrm{d}t \lesssim C_0|t|^{-\frac{3}{k-2}}. \tag{5.35}$$

Again from (5.23) and (5.25) we have $|\lambda'(t) - b(t)| \lesssim C_0|t|^{-\frac{k+1}{k-2}}$. Multiplying by $b'(t) = \frac{k}{2}\cdot\big(\frac{2\kappa}{k-2}\big)^k\cdot\frac{\lambda(t)^{k-1}}{\mu(t)^k} \sim |t|^{-\frac{2k-2}{k-2}}$, cf. (5.9) and (5.8), we obtain $\big|\frac{\mathrm{d}}{\mathrm{d}t}\big(b(t)^2 - \big(\frac{2\kappa}{k-2}\big)^k\frac{\lambda(t)^k}{\mu(t)^k}\big)\big| \lesssim C_0|t|^{-\frac{3k-1}{k-2}}$. Since $b(T) = \big(\frac{2\kappa}{k-2}\big)^{\frac{k}{2}}\cdot\lambda(T)$ and $\mu(T) = 1$, this yields $\big|b(t)^2 - \big(\frac{2\kappa}{k-2}\big)^k\frac{\lambda(t)^k}{\mu(t)^k}\big| \lesssim C_0|t|^{-\frac{2k+1}{k-2}}$. But $b(t) + \big(\frac{2\kappa}{k-2}\big)^{\frac{k}{2}}\frac{\lambda(t)^{\frac{k}{2}}}{\mu(t)^{\frac{k}{2}}} \sim |t|^{\frac{k}{k-2}}$, see (2.7) and (2.9), hence

$$\big|b(t) - \big(\frac{2\kappa}{k-2}\big)^{\frac{k}{2}}\frac{\lambda(t)^{\frac{k}{2}}}{\mu(t)^{\frac{k}{2}}}\big| \lesssim C_0|t|^{-\frac{k+1}{k-2}}. \tag{5.36}$$

Bound (5.35) implies that $\big|\frac{\lambda(t)^{\frac{k}{2}}}{\mu(t)^{\frac{k}{2}}} - \lambda(t)^{\frac{k}{2}}\big| \ll |t|^{-\frac{k+1}{k-2}}$, thus (5.36) yields $|\lambda'(t) - \big(\frac{2\kappa}{k-2}\big)^{\frac{k}{2}}\lambda(t)^{\frac{k}{2}}| \lesssim C_0|t|^{-\frac{k+1}{k-2}}$. Integrating this differential inequality is standard. Dividing by $\lambda(t)^{\frac{k}{2}} \sim |t|^{-\frac{k}{k-2}}$ we get

$$\big|(\lambda^{-\frac{k-2}{2}})' + \frac{k-2}{2}\big(\frac{2\kappa}{k-2}\big)^{\frac{k}{2}}\big| = \big|(\lambda^{-\frac{k-2}{2}})' + \big(\frac{2\kappa}{k-2}\big)^{\frac{k-2}{2}}\kappa\big| \lesssim C_0|t|^{-\frac{1}{k-2}}.$$

Using $\lambda(T) = \frac{k-2}{2\kappa}(\kappa|T|)^{-\frac{2}{k-2}}$ we obtain $\big|\lambda(t)^{-\frac{k-2}{2}} - \big(\frac{2\kappa}{k-2}\big)^{\frac{k-2}{2}}\kappa|t|\big| \lesssim C_0|t|^{\frac{k-3}{k-2}}$, from which (5.27) follows.

We turn to the proof of (5.26). From (5.10) the initial data at $t = T$ satisfy $\|\boldsymbol{g}(T)\|_{\mathcal{E}} \lesssim |t|^{-\frac{2k}{k-2}} \ll |t|^{-\frac{k+1}{k-2}}$, thus $H(T) \lesssim |t|^{-\frac{2k+2}{k-2}}$. If C_0 is large enough, then integrating (5.34) we get $H(t) \leq c\cdot C_0^2|t|^{-\frac{2k+2}{k-2}}$, with a small constant c. Eventually C_0 if necessary and using the coercivity of H, we obtain (5.26). □

A Cauchy theory

A.1 Persistence of regularity

Proposition A.1. *Let* $\boldsymbol{u} : (T_-, T_+) \to \mathcal{E}$ *be the solution of* (1.1) *with the initial condition* $\boldsymbol{u}(t_0) = \boldsymbol{u}_0$. *If* $\boldsymbol{u}_0 \in X^1 \times H^1$, *then* $\boldsymbol{u} \in C((T_-, T_+); X^1 \times H^1) \cap C^1((T_-, T_+); \mathcal{E})$. □

The proof is classical, see [9, Chapter 5] for more general results in the case of the nonlinear Schrödinger equation. Analogous results hold in the case of equations (1.4) and (1.3).

A.2 Profile decomposition and consequences

For details about the nonlinear profile decomposition for the critical wave equation we refer to [3] (the defocusing case), [23] (dimension $N = 3$) and [81] (any dimension). For the reader's convenience we recall the following result [81, Proposition 2.3]. We denote $S(I)$ the Strichartz norm on the interval I.

Proposition A.2. *Let $\boldsymbol{u}_{0,n}$ be a bounded sequence in $\mathcal{E}$ admitting a profile decomposition with profiles $\boldsymbol{U}^j_{\mathrm{LIN}}$ and parameters $\lambda_{j,n}$, $t_{j,n}$. Let $\boldsymbol{U}^j$ be the corresponding nonlinear profiles and let θ_n be a sequence of positive numbers. Assume*

$$\forall j \geq 1, n \geq 1, \ \frac{\theta_n - t_{j,n}}{\lambda_{j,n}} < T_+(U^j) \quad \text{and} \quad \limsup_{n\to+\infty} \|U^j\|_{S\left(\frac{-t_{j,n}}{\lambda_{j,n}}, \frac{\theta_n - t_{j,n}}{\lambda_{j,n}}\right)} < +\infty.$$

Let $\boldsymbol{u}_n$ be the solution of (1.1) *with the initial data $\boldsymbol{u}_n(0) = \boldsymbol{u}_{0,n}$. Then for n sufficiently large $\boldsymbol{u}_n$ is defined on $[0, \theta_n]$ and*

$$\boldsymbol{u}_n(t) = \sum_{j=1}^{J} \boldsymbol{U}^j_n(t) + \boldsymbol{w}^J_n(t) + \boldsymbol{r}^J_n(t), \qquad \text{for all } J \in \mathbb{N} \text{ and } t \in [0, \theta_n],$$

with $\lim_{J\to+\infty} \limsup_{n\to+\infty} \sup_{t\in[0,\theta_n]} \|\boldsymbol{r}^J_n\|_{\mathcal{E}} = 0$. An analogous statement holds for $\theta \leq 0$.

For a corresponding result for the Yang-Mills equation and the equivariant wave maps, see [42].

Corollary A.3. *There exists a constant $\eta > 0$ such that the following holds. Let $\boldsymbol{u} : [t_0, T_+) \to \mathcal{E}$ be a maximal solution of* (1.1) *with $T_+ < +\infty$. Then for any compact set $K \subset \mathcal{E}$ there exists $\tau < T_+$ such that $\mathrm{dist}(\boldsymbol{u}(t), K) > \eta$ for $t \in [\tau, T_+)$.*

Proof. Suppose for the sake of contradiction that there exists a sequence $t_n \to T_+$ such that for $\boldsymbol{u}_{0,n} := \boldsymbol{u}(t_n)$ there holds $\mathrm{dist}(\boldsymbol{u}_{0,n}, K) \leq \eta$, hence $\boldsymbol{u}_{0,n} = \boldsymbol{k}_n + \boldsymbol{b}_n$ with $\boldsymbol{k}_n \in K$ and $\|\boldsymbol{b}_n\|_{\mathcal{E}} \leq \eta$. By taking a subsequence we can assume that $\boldsymbol{k}_n \to \boldsymbol{k}_0 \in K$ and that the sequence $\boldsymbol{b}_n$ admits a profile decomposition, in particular $\boldsymbol{b}_n \rightharpoonup \boldsymbol{b}_0 \in \mathcal{E}$. This implies that $\boldsymbol{u}_{0,n}$ admits a profile decomposition with the first profile $\boldsymbol{k}_0 + \boldsymbol{b}_0$ and all the other profiles small in the energy norm. Proposition A.2 leads to a contradiction. □

Corollary A.4. *There exists a constant $\eta > 0$ such that the following holds. Let $K \subset \mathcal{E}$ be a compact set and let $\boldsymbol{u}_n : [T_1, T_2] \to \mathcal{E}$ be a sequence of solutions of* (1.1) *such that*

$$\mathrm{dist}(\boldsymbol{u}_n(t), K) \leq \eta, \qquad \text{for all } n \in \mathbb{N} \text{ and } t \in [T_1, T_2].$$

Suppose that $\boldsymbol{u}_n(T_1) \rightharpoonup \boldsymbol{u}_0 \in \mathcal{E}$. Then the solution $\boldsymbol{u}(t)$ of (1.1) *with the initial condition $\boldsymbol{u}(T_1) = \boldsymbol{u}_0$ is defined for $t \in [T_1, T_2]$ and*

$$\boldsymbol{u}_n(t) \rightharpoonup \boldsymbol{u}(t), \qquad \text{for all } t \in [T_1, T_2]. \tag{A.1}$$

Proof. It suffices to prove (A.1) for a subsequence of any subsequence. Hence, we may assume that $\boldsymbol{u}_{0,n} := \boldsymbol{u}_n(T_1)$ admits a profile decomposition. As in the previous proof, we show that all the profiles except for the weak limit of $\boldsymbol{u}_{0,n}$ are small in the energy norm. Proposition A.2 implies (A.1). □

Remark A.5. Corollaries A.3 and A.4 hold for the Yang-Mills and wave map equations, with the same proofs.

Remark A.6. An important point of both results is that η is independent of K. Corollary A.3 states that a blow-up cannot happen at a small distance from a compact set. Corollary A.4 establishes sequential weak continuity of the flow in a neighbourhood of any compact set. Without this additional condition weak continuity is expected to fail, a counterexample being provided by type II blow-up solutions.

Remark A.7. Corollaries A.3 and A.4 are crucial ingredients of the arguments in Subsection 3.6. Using the nonlinear profile decomposition of Bahouri and Gérard [3] is nowadays a well-established method of attacking this type of questions in critical spaces. Note that [3] gave the first proof of the sequential weak continuity of the flow for the defocusing energy-critical wave equation.

Chapter 3

Bounds on the speed of type II blow-up for the energy critical wave equation in the radial case

Abstract

We consider the focusing energy-critical wave equation in space dimension $N \in \{3, 4, 5\}$ for radial data. We study type II blow-up solutions which concentrate one bubble of energy. It is known that such solutions decompose in the energy space as a sum of the bubble and an asymptotic profile. We prove bounds on the blow-up speed in the case when the asymptotic profile is sufficiently regular. These bounds are optimal in dimension $N = 5$. We also prove that if the asymptotic profile is sufficiently regular, then it cannot be strictly negative at the origin.

1 Introduction

1.1 Setting of the problem

Let $N \in \{3,4,5\}$ be the dimension of the space. For $\boldsymbol{u}_0 = (u_0, \dot u_0) \in \mathcal{E} := \dot H^1(\mathbb{R}^N) \times L^2(\mathbb{R}^N)$, define the *energy functional*

$$E(\boldsymbol{u}_0) = \int \frac{1}{2}|\dot u_0|^2 + \frac{1}{2}|\nabla u_0|^2 - F(u_0)\,\mathrm{d}x,$$

where $F(u_0) := \frac{N-2}{2N}|u_0|^{\frac{2N}{N-2}}$. Note that $E(\boldsymbol{u}_0)$ is well-defined due to the Sobolev Embedding Theorem. The differential of E is $\mathrm{D}E(\boldsymbol{u}_0) = (-\Delta u_0 - f(u_0), \dot u_0)$, where $f(u_0) = |u_0|^{\frac{4}{N-2}}u_0$.

We consider the Cauchy problem for the energy critical wave equation:

$$\begin{cases} \partial_t \boldsymbol{u}(t) = J \circ \mathrm{D}E(\boldsymbol{u}(t)), \\ \boldsymbol{u}(t_0) = \boldsymbol{u}_0 \in \mathcal{E}. \end{cases} \tag{NLW}$$

Here, $J := \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix}$ is the natural symplectic structure. This equation is often written in the form

$$\partial_{tt} u = \Delta u + f(u).$$

Equation (NLW) is locally well-posed in the space $\mathcal{E}$, see for example [32] and [84] (the defocusing case), as well as a complete review of the Cauchy theory in [47]. In particular, for any initial data $\boldsymbol{u}_0 \in \mathcal{E}$ there exists a maximal time of existence (T_-, T_+), $-\infty \le T_- < t_0 < T_+ \le +\infty$, and a unique solution $\boldsymbol{u} \in C((T_-, T_+); \mathcal{E})$. In addition, the energy E is a conservation law. In this paper we always assume that the initial data is radially symmetric. This symmetry is preserved by the flow.

For functions $v \in \dot H^1$, $\dot v \in L^2$, $\boldsymbol{v} = (v, \dot v) \in \mathcal{E}$ and $\lambda > 0$, we denote

$$v_\lambda(x) := \frac{1}{\lambda^{(N-2)/2}} v\Big(\frac{x}{\lambda}\Big), \qquad \dot v_{\underline{\lambda}}(x) := \frac{1}{\lambda^{N/2}} \dot v\Big(\frac{x}{\lambda}\Big), \qquad \boldsymbol{v}_\lambda(x) := \big(v_\lambda, \dot v_{\underline{\lambda}}\big).$$

A change of variables shows that

$$E\big((\boldsymbol{u}_0)_\lambda\big) = E(\boldsymbol{u}_0).$$

Equation (NLW) is invariant under the same scaling. If $\boldsymbol{u} = (u, \dot u)$ is a solution of (NLW) and $\lambda > 0$, then $t \mapsto \boldsymbol{u}\big((t - t_0)/\lambda\big)_\lambda$ is also a solution with initial data $(\boldsymbol{u}_0)_\lambda$ at time $t = 0$. This is why equation (NLW) is called *energy-critical.*

A fundamental object in the study of (NLW) is the family of stationary solutions $(u, \partial_t u) = \pm \boldsymbol{W}_\lambda = (\pm W_\lambda, 0)$, where

$$W(x) = \Big(1 + \frac{|x|^2}{N(N-2)}\Big)^{-(N-2)/2}.$$

The functions W_λ are called *ground states.*

In general the energy E does not control the norm $\|\cdot\|_{\mathcal{E}}$, and indeed this norm can tend to $+\infty$ in finite time, which is referred to as *type I blow-up.* In odd space dimensions and for superconformal nonlinearities (which includes the energy-critical case) Donninger and Schörkhuber [21], [22] described large sets of initial data leading to this kind of blow-up.

It can also happen that in finite time the solution leaves every compact set of $\mathcal{E}$, the norm $\|\cdot\|_{\mathcal{E}}$ staying bounded, which is referred to as *type II blow-up.* In dimension $N = 3$ in the radial case one of the consequences of the classification result of Duyckaerts, Kenig and Merle [26] is that any blow-up solution is either of type I or of type II. This is unknown in other cases.

A particular type of type II blow-up occurs when the solution $\boldsymbol{u}(t)$ stays close to the family of ground states $\boldsymbol{W}_\lambda$ and $\lambda \to 0$. In this situation we call $\boldsymbol{W}_\lambda$ the *bubble of energy* and we say that $\boldsymbol{u}(t)$ blows up by concentration of one bubble of energy. We have the following fundamental result proved first by Duyckaerts, Kenig and Merle [23] for $N = 3$, by the same authors [25] for $N = 5$ and by Côte, Kenig, Lawrie and Schlag [15] for $N = 4$:

Theorem ([23], [25], [15]). *Let $\boldsymbol{u}(t)$ be a radial solution of* (NLW) *which blows up at $t = T_+$ by concentration of one bubble of energy. Then there exist $\boldsymbol{u}_0^* \in \mathcal{E}$ and $\lambda \in C([t_0, T_+), (0, +\infty))$ such that*

$$\lim_{t\to T_+} \|\boldsymbol{u}(t) - \boldsymbol{W}_{\lambda(t)} - \boldsymbol{u}_0^*\|_{\mathcal{E}} = 0, \qquad \lim_{t\to T_+} (T_+ - t)^{-1}\lambda(t) = 0. \tag{1.1}$$

□

In this context the function $\boldsymbol{u}_0^*$ is called the *asymptotic profile*. Note that in [25] a more general, non-radial version of the above theorem was proved for $N \in \{3, 5\}$.

Solutions verifying (1.1) were first constructed in dimension $N = 3$ by Krieger, Schlag and Tataru [53], who obtained all possible polynomial blow-up rates $\lambda(t) \sim (T_+ - t)^{1+\nu}$, $\nu > 0$. For $N = 4$ smooth solutions blowing up at a particular rate were constructed by Hillairet and Raphaël [36]. For $N = 5$ the author proved in [40] that for any radially symmetric asymptotic profile $\boldsymbol{u}_0^* \in H^4 \times H^3$ such that $u_0^*(0) > 0$, there exists a solution $\boldsymbol{u}(t)$ such that (1.1) holds. For these solutions the concentration speed of the bubble is

$$\lambda(t) \sim u_0^*(0)^2 (T_+ - t)^4. \tag{1.2}$$

In the same article, solutions with blow-up rate $(T_+ - t)^{1+\nu}$ for $\nu > 8$ were constructed, with ν explicitely related to the asymptotic behaviour of $\boldsymbol{u}_0^*$ at $x = 0$.

1.2 Statement of the results

In the present paper we continue the investigation of the relationship between the behaviour of $\boldsymbol{u}_0^*$ at $x = 0$ and possible blow-up speeds, still in the special case when the asymptotic profile $\boldsymbol{u}_0^*$ is sufficiently regular. We prove the following result.

Theorem 1. *Let $N \in \{3, 4, 5\}$ and $s > \frac{N-2}{2}$, $s \geq 1$. Let $\boldsymbol{u}_0^* = (u_0^*, \dot{u}_0^*) \in H^{s+1} \times H^s$ be a radial function. Suppose that $\boldsymbol{u}$ is a radial solution of* (NLW) *such that*

$$\lim_{t\to T_+} \|\boldsymbol{u}(t) - \boldsymbol{W}_{\lambda(t)} - \boldsymbol{u}_0^*\|_{\mathcal{E}} = 0, \qquad \lim_{t\to T_+} \lambda(t) = 0, \qquad T_+ < +\infty. \tag{1.3}$$

There exists a constant $C > 0$ depending on $\boldsymbol{u}_0^*$ such that:

- *if $N \in \{4, 5\}$, then for $T_+ - t$ sufficiently small there holds*

$$\lambda(t) \leq C(T_+ - t)^{\frac{4}{6-N}}. \tag{1.4}$$

- *if $N = 3$, then there exists a sequence $t_n \to T_+$ such that*

$$\lambda(t_n) \leq C(T_+ - t_n)^{\frac{4}{6-N}}. \tag{1.5}$$

Remark 1.1. Let $\boldsymbol{u}^* = (u^*, \dot{u}^*)$ be the solution of (NLW) such that $\boldsymbol{u}^*(T_+) = \boldsymbol{u}_0^*$ and suppose that $0 \in \operatorname{supp} \boldsymbol{u}_0^*$. We will prove that there exists a universal constant C_0 such that in the above theorem one can take

$$C = C_0 \|u^*\|_{L^\infty((T_+-\rho, T_+)\times B(0,\rho))}^{\frac{2}{6-N}},$$

where $\rho > 0$ is arbitrary and $B(0,\rho)$ is the ball of centre 0 and radius ρ in $\mathbb{R}^N$. Notice that $u^* \in L^\infty((T_+ - \rho, T_+) \times \mathbb{R}^N)$ by Appendix A and the Sobolev Embedding Theorem.

If $0 \notin \operatorname{supp} \boldsymbol{u}_0^*$, then blow-up cannot occur, as follows from the classification of solutions of (NLW) at energy level $E(\boldsymbol{W})$ by Duyckaerts and Merle [27].

Remark 1.2. In the case $N = 3$ we will prove that for $T_+ - t$ small enough there holds

$$\int_t^{T_+} \frac{\mathrm{d}\tau}{\sqrt{\lambda(\tau)}} \geq \frac{3}{\sqrt{C}}(T_+ - t)^{\frac{1}{3}}, \tag{1.6}$$

which immediately implies (1.5).

If we assume that $\boldsymbol{u}^* \in H^3 \times H^2$, then (1.4) holds also in the case $N = 3$, see Remark 2.13. I believe that the proof of (1.5) given here could be adapted to cover the case $1 > s > \frac{1}{2}$.

Remark 1.3. In dimension $N = 5$ the bound (1.4) is optimal, see (1.2). It is not clear if the bounds are optimal for $N \in \{3, 4\}$, due to slow decay of the bubble.

Remark 1.4. A natural problem is to determine sharp bounds for the blow-up speed in the case of less regular $\boldsymbol{u}_0^*$. The method used in this paper allows to obtain some bounds for example in the case $1 \leq s < \frac{3}{2}$ in dimension $N = 5$, but they are not optimal and I will not pursue this direction here.

In the case $u_0^*(0) = 0$ one could obtain various bounds depending on the asymptotics of $\boldsymbol{u}_0^*$ at $x = 0$, but this will not be considered in the present paper. Along the same line, one can ask if the sign of $u_0^*(0)$ is relevant in the case when $u_0^*(0) \neq 0$. It turns out that it is, but unfortunately our method requires the additional assumption $\boldsymbol{u}_0^* \in H^3 \times H^2$:

Theorem 2. *Let $N \in \{3, 4, 5\}$. Let $\boldsymbol{u}_0^* = (u_0^*, \dot{u}_0^*) \in H^3 \times H^2$ be a radial function such that*

$$u_0^*(0) < 0.$$

There exist no radial solutions of (NLW) *such that*

$$\lim_{t\to T_+} \|\boldsymbol{u}(t) - \boldsymbol{W}_{\lambda(t)} - \boldsymbol{u}_0^*\|_{\mathcal{E}} = 0, \qquad \lim_{t\to T_+} \lambda(t) = 0, \qquad T_+ < +\infty.$$

Remark 1.5. I expect that Theorems 1 and 2 could be proved by similar methods without the assumption of $\boldsymbol{u}_0^*$ being radial.

1.3 Related results

The problem of existence of an asymptotic profile at blow-up might be seen as a version of the classical question of asymptotic stability of solitons in the case when finite-time blow-up occurs. Decompositions of type (1.1) in suitable topologies are believed to hold for many models, but establishing this rigourously is a challenging problem. Historically, the study of finite type blow-up in the Hamiltonian setting received the most attention probably in the case of nonlinear Schrödinger equations (NLS). For the mass-critical NLS the conformal invariance leads to explicit blow-up solutions $S(t)$ with the asymptotic profile $u^* \equiv 0$. Bourgain and Wang [5] constructed examples of blow-up solutions with u^* regular and non-zero, the speed of blow-up however being the same as for $S(t)$. This is not a coincidence, as shown by a classification result of Merle and Raphaël [65].

For the critical gKdV equation Martel, Merle and Raphaël [59] proved that if the initial data decays sufficiently fast, then there is only one possible blow-up speed, given by the minimal mass blow-up solution. However, without the decay assumption other blow-up speeds are possible, as shown by the same authors in [60].

These are the main two examples of the heuristic principle that the size of the interaction of the bubble with the rest of the solution influences or even determines the speed of blow-up. In the present paper we try to investigate this phenomenon in the energy-critical setting. The same problem could be considered for other energy-critical models for which bubble concentration phenomenon has been observed (see results of Krieger, Schlag and Tataru [52], Ortoleva and Perelman [74], Perelman [76], Merle, Rodnianski and Raphaël [66], Schweyer [82]). It seems that the question of relationship between the asymptotic profile and the speed of blow-up has not been addressed.

Finally, let us mention that the problem of understanding the possible blow-up speeds is not limited to type II blow-up for critical equations, see for instance Merle and Zaag [67], [68] for the subconformal and conformal NLW, and Giga and Kohn [31], Mizoguchi [69], Matano and Merle [63] for the semilinear heat equation.

1.4 Outline of the proof

Our proofs of Theorems 1 and 2 are based on the following computation that we present here formally.

Let $\boldsymbol{u} : [t_0; T_+) \to \mathcal{E}$ be a solution of (NLW) which satisfies (1.1). At blow-up time, the energy of the bubble is completely decoupled from the energy of the asymptotic profile, hence

$$E(\boldsymbol{u}) = E(\boldsymbol{u}_0^*) + E(\boldsymbol{W}_\lambda) = E(\boldsymbol{u}_0^*) + E(\boldsymbol{W}). \tag{1.7}$$

Let $\boldsymbol{u}^*$ be the solution of (NLW) with the initial data $\boldsymbol{u}^*(T_+) = \boldsymbol{u}_0^*$. Decompose $\boldsymbol{u}(t) = \boldsymbol{W}_{\lambda(t)} + \boldsymbol{u}^*(t) + \boldsymbol{g}(t)$ (in fact we use a suitable localisation of $\boldsymbol{W}_\lambda$; we will ignore here this technical point). The *modulation parameter* λ is determined by a suitable orthogonality condition, and a standard procedure shows that $|\lambda'(t)| \lesssim \|\boldsymbol{g}(t)\|_{\mathcal{E}}$.

From the Taylor formula we obtain

$$E(\boldsymbol{u}) = E(\boldsymbol{u}^* + \boldsymbol{W}_\lambda) + \langle \mathrm{D}E(\boldsymbol{u}^* + \boldsymbol{W}_\lambda), \boldsymbol{g}\rangle + \frac{1}{2}\langle \mathrm{D}^2 E(\boldsymbol{u}^* + \boldsymbol{W}_\lambda)\boldsymbol{g}, \boldsymbol{g}\rangle + O(\|\boldsymbol{g}\|_{\mathcal{E}}^3).$$

Step 1. An explicit key computation shows that

$$E(\boldsymbol{u}^* + \boldsymbol{W}_\lambda) - E(\boldsymbol{u}^*) - E(\boldsymbol{W}) \gtrsim -u_0^*(0)\lambda^{\frac{N-2}{2}}.$$

It is clear that the sign of $u_0^*(0)$ is decisive.

Step 2. Near blow-up time $\boldsymbol{u}^*$ weakly interacts with $\boldsymbol{W}_\lambda$ and $\mathrm{D}E(\boldsymbol{W}_\lambda) = 0$. This allows to replace $\langle \mathrm{D}E(\boldsymbol{u}^* + \boldsymbol{W}_\lambda), \boldsymbol{g}\rangle$ by $\langle \mathrm{D}E(\boldsymbol{u}^*), \boldsymbol{g}\rangle$. Using the Hamiltonian structure it is seen that this quantity is, at first order in $\boldsymbol{g}$, a conservation law. Estimating some error terms we conclude that this term can be neglected.

Step 3. Let us suppose for a moment that $\mathrm{D}^2E(\boldsymbol{W})$ is a coercive functional in the sense that $\langle \mathrm{D}^2E(\boldsymbol{u}^*+\boldsymbol{W}_\lambda)\boldsymbol{g},\boldsymbol{g}\rangle \gtrsim \|\boldsymbol{g}\|_{\mathcal{E}}^2$. Using (1.7) and the two preceding steps we find $|\lambda'|^2 \lesssim \|\boldsymbol{g}\|_{\mathcal{E}}^2 \lesssim u_0^*(0)\lambda^{\frac{N-2}{2}}$. In the case $u_0^*(0) < 0$ this is contradictory, and in the case $u_0^*(0) > 0$ the conclusion follows by integrating the differential inequality for λ.

Strictly speaking, $\mathrm{D}^2E(\boldsymbol{W})$ is not a coercive functional, and much of the proof is devoted to controlling the negative directions, which are related to the eigendirections of the flow linearized around $\boldsymbol{W}$. Clarifying the second step above is another major technical difficulty of this paper.

1.5 Notation

We introduce the inifinitesimal generators of scale change

$$\Lambda_s := \big(\frac{N}{2}-s\big)+x\cdot\nabla.$$

For $s=1$ we omit the subscript and write $\Lambda=\Lambda_1$. We denote $\Lambda_{\mathcal{E}}$, $\Lambda_{\mathcal{F}}$ and $\Lambda_{\mathcal{E}^*}$ the inifinitesimal generators of the scaling which is critical for a given norm, that is

$$\Lambda_{\mathcal{E}} = (\Lambda,\Lambda_0),\quad \Lambda_{\mathcal{E}^*} = (\Lambda_{-1},\Lambda_0).$$

The dimension of the space will be denoted N. The domain of the function spaces is always $\mathbb{R}^N$. We introduce the following notation for some frequently used function spaces: $X^s := \dot H^{s+1}\cap\dot H^1$ for $s\geq 0$, $\mathcal{E} := \dot H^1\times L^2$, $\mathcal{F} := L^2\times\dot H^{-1}$. The bracket $\langle\cdot,\cdot\rangle$ denotes the distributional pairing and the scalar product in the spaces L^2, $L^2\times L^2$. Notice that $\mathcal{E}^*\simeq\dot H^{-1}\times L^2$ through the natural isomorphism induced by $\langle\cdot,\cdot\rangle$.

For a function space $\mathcal{A}$, $O_{\mathcal{A}}(m)$ denotes any $a\in\mathcal{A}$ such that $\|a\|_{\mathcal{A}}\leq Cm$ for some constant $C>0$. For positive quantities m_1 and m_2 we write $m_1\lesssim m_2$ for $m_1 = O(m_2)$ and $m_1\sim m_2$ for $m_1\lesssim m_2\lesssim m_1$. We denote $B_{\mathcal{A}}(x_0,\delta)$ the open ball of center x_0 and radius δ in the space $\mathcal{A}$. If $\mathcal{A}$ is not specified, it means that $\mathcal{A}=\mathbb{R}$.

2 The proofs

2.1 Properties of the linearized operator

Linearizing $-\Delta u - f(u)$ around W, $u = W+g$, we obtain a Schrödinger operator

$$Lg = (-\Delta - f'(W))g.$$

Notice that $L(\Lambda W) = \frac{\mathrm{d}}{\mathrm{d}\lambda}\big|_{\lambda=1}\big(-\Delta W_\lambda - f(W_\lambda)\big) = 0$. It is known that L has exactly one strictly negative simple eigenvalue which we denote $-\nu^2$ (we take $\nu>0$). We denote the corresponding positive eigenfunction $\mathcal{Y}$, normalized so that $\|\mathcal{Y}\|_{L^2}=1$. By elliptic regularity $\mathcal{Y}$ is smooth and by Agmon estimates it decays exponentially. Self-adjointness of L implies that

$$\langle\mathcal{Y},\Lambda W\rangle = 0. \tag{2.1}$$

We define

$$\mathcal{Y}^- := \big(\frac{1}{\nu}\mathcal{Y},-\mathcal{Y}\big),\qquad \mathcal{Y}^+ := \big(\frac{1}{\nu}\mathcal{Y},\mathcal{Y}\big),\qquad \alpha^- := \frac12(\nu\mathcal{Y},-\mathcal{Y}),\qquad \alpha^+ := \frac12(\nu\mathcal{Y},\mathcal{Y}).$$

We have $J \circ \mathrm{D}^2 E(\boldsymbol{W}) = \begin{pmatrix} 0 & \mathrm{Id} \\ -L & 0 \end{pmatrix}$. A short computation shows that

$$J \circ \mathrm{D}^2 E(\boldsymbol{W})\mathcal{Y}^- = -\nu\mathcal{Y}^-, \qquad J \circ \mathrm{D}^2 E(\boldsymbol{W})\mathcal{Y}^+ = \nu\mathcal{Y}^+$$

and

$$\langle \alpha^-, J \circ \mathrm{D}^2 E(\boldsymbol{W})\boldsymbol{g}\rangle = -\nu\langle \alpha^-, \boldsymbol{g}\rangle, \qquad \langle \alpha^+, J \circ \mathrm{D}^2 E(\boldsymbol{W})\boldsymbol{g}\rangle = \nu\langle \alpha^+, \boldsymbol{g}\rangle, \qquad \forall \boldsymbol{g} \in \mathcal{E}. \tag{2.2}$$

Notice that $\langle \alpha^-, \mathcal{Y}^-\rangle = \langle \alpha^+, \mathcal{Y}^+\rangle = 1$ and $\langle \alpha^-, \mathcal{Y}^+\rangle = \langle \alpha^+, \mathcal{Y}^-\rangle = 0$.

The rescaled versions of these objects are

$$\mathcal{Y}_\lambda^- := \big(\frac{1}{\nu}\mathcal{Y}_\lambda, -\mathcal{Y}_{\underline{\lambda}}\big), \qquad \mathcal{Y}_\lambda^+ := \big(\frac{1}{\nu}\mathcal{Y}_\lambda, \mathcal{Y}_{\underline{\lambda}}\big), \qquad \alpha_\lambda^- := \frac{1}{2}\big(\frac{\nu}{\lambda}\mathcal{Y}_{\underline{\lambda}}, -\mathcal{Y}_{\underline{\lambda}}\big), \qquad \alpha_\lambda^+ := \frac{1}{2}\big(\frac{\nu}{\lambda}\mathcal{Y}_{\underline{\lambda}}, \mathcal{Y}_{\underline{\lambda}}\big).$$

The scaling is chosen so that $\langle \alpha_\lambda^-, \mathcal{Y}_\lambda^-\rangle = \langle \alpha_\lambda^+, \mathcal{Y}_\lambda^+\rangle = 1$. We have

$$J \circ \mathrm{D}^2 E(\boldsymbol{W}_\lambda)\mathcal{Y}_\lambda^- = -\frac{\nu}{\lambda}\mathcal{Y}_\lambda^-, \qquad J \circ \mathrm{D}^2 E(\boldsymbol{W}_\lambda)\mathcal{Y}_\lambda^+ = \frac{\nu}{\lambda}\mathcal{Y}_\lambda^+$$

and

$$\langle \alpha_\lambda^-, J \circ \mathrm{D}^2 E(\boldsymbol{W}_\lambda)\boldsymbol{g}\rangle = -\frac{\nu}{\lambda}\langle \alpha_\lambda^-, \boldsymbol{g}\rangle, \qquad \langle \alpha_\lambda^+, J \circ \mathrm{D}^2 E(\boldsymbol{W}_\lambda)\boldsymbol{g}\rangle = \frac{\nu}{\lambda}\langle \alpha_\lambda^+, \boldsymbol{g}\rangle, \qquad \forall \boldsymbol{g} \in \mathcal{E}.$$

Let $\mathcal{Z}$ be a C_0^∞ function such that

$$\langle \mathcal{Z}, \Lambda W\rangle > 0, \qquad \langle \mathcal{Z}, \mathcal{Y}\rangle = 0$$

(the first condition is the essential one and the second allows to simplify some computations). We recall the following result.

Proposition 2.1 ([40, Lemma 6.1], [27, Proposition 5.5])**.** *There exists a constant $c_L > 0$ such that*

$$v \in \dot H^1 \textit{ radial}, \quad \langle \mathcal{Y}, v\rangle = \langle \mathcal{Z}, v\rangle = 0 \quad \Rightarrow \quad \frac{1}{2}\langle v, Lv\rangle \geq c_L\|v\|_{\dot H^1}^2.$$

□

Lemma 2.2. *There exists a constant $c > 0$ such that if $\|\boldsymbol{V} - \boldsymbol{W}_\lambda\|_\mathcal{E} < c$, then for all $\boldsymbol{g} \in \mathcal{E}$ such that $\langle \mathcal{Z}_{\underline{\lambda}}, g\rangle = 0$ there holds*

$$\frac{1}{2}\langle \mathrm{D}^2 E(\boldsymbol{V})\boldsymbol{g}, \boldsymbol{g}\rangle + 2\big(\langle \alpha_\lambda^-, \boldsymbol{g}\rangle^2 + \langle \alpha_\lambda^+, \boldsymbol{g}\rangle^2\big) \gtrsim \|\boldsymbol{g}\|_\mathcal{E}^2.$$

Proof. We have

$$\langle \mathrm{D}^2 E(\boldsymbol{V})\boldsymbol{g}, \boldsymbol{g}\rangle = \langle \mathrm{D}^2 E(\boldsymbol{W}_\lambda)\boldsymbol{g}, \boldsymbol{g}\rangle + \int \big(f'(V) - f'(W_\lambda)\big)|g|^2\,\mathrm{d}x.$$

By Hölder, the last integral is $\lesssim c\|\boldsymbol{g}\|_\mathcal{E}^2$, hence it suffices to prove the lemma with $\boldsymbol{V} = \boldsymbol{W}_\lambda$. Without loss of generality we can assume that $\lambda = 1$. We will show the following stronger inequality:

$$\frac{1}{2}\langle \mathrm{D}^2 E(\boldsymbol{W})\boldsymbol{g}, \boldsymbol{g}\rangle + 2\langle \alpha^-, \boldsymbol{g}\rangle \cdot \langle \alpha^+, \boldsymbol{g}\rangle \geq c_L\|\boldsymbol{g} - \langle \alpha^-, \boldsymbol{g}\rangle\mathcal{Y}^- - \langle \alpha^+, \boldsymbol{g}\rangle\mathcal{Y}^+\|_\mathcal{E}^2. \tag{2.3}$$

Let $a^- = \langle \alpha^-, \boldsymbol{g}\rangle$, $a^+ = \langle \alpha^+, \boldsymbol{g}\rangle$ and decompose $\boldsymbol{g} = a^-\mathcal{Y}^- + a^+\mathcal{Y}^+ + \boldsymbol{k}$, so that $\langle \alpha^-, \boldsymbol{k}\rangle = \langle \alpha^+, \boldsymbol{k}\rangle = 0$. From $\langle \mathcal{Z}, \mathcal{Y}\rangle = 0$ we deduce $\langle \mathcal{Z}, k\rangle = 0$. We have $g = \frac{a^- + a^+}{\nu}\mathcal{Y} + k$ and $\dot g = (-a^- + a^+)\mathcal{Y} + \dot k$, hence

$$\begin{aligned}
\frac{1}{2}\langle \mathrm{D}^2 E(\boldsymbol{W})\boldsymbol{g}, \boldsymbol{g}\rangle &= \frac{1}{2}\Big\langle \frac{a^- + a^+}{\nu}\mathcal{Y} + k, -(a^- + a^+)\nu\mathcal{Y} + Lk\Big\rangle \\
&+ \frac{1}{2}\big\langle (-a^- + a^+)\mathcal{Y} + \dot k, (-a^- + a^+)\mathcal{Y} + \dot k\big\rangle \\
&= -\frac{1}{2}(a^- + a^+)^2\langle \mathcal{Y}, \mathcal{Y}\rangle - (a^- + a^+)\nu\langle \mathcal{Y}, k\rangle + \frac{1}{2}\langle k, Lk\rangle \\
&+ \frac{1}{2}(-a^- + a^+)^2\langle \mathcal{Y}, \mathcal{Y}\rangle + (-a^- + a^+)\langle \mathcal{Y}, \dot k\rangle + \frac{1}{2}\langle \dot k, \dot k\rangle \\
&= -2a^- a^+\langle \mathcal{Y}, \mathcal{Y}\rangle - 2a^-\langle \alpha^+, \boldsymbol{k}\rangle - 2a^+\langle \alpha^-, \boldsymbol{k}\rangle + \frac{1}{2}\big(\langle k, Lk\rangle + \langle \dot k, \dot k\rangle\big) \\
&= -2a^- a^+ + \frac{1}{2}\langle \mathrm{D}^2 E(\boldsymbol{W})\boldsymbol{k}, \boldsymbol{k}\rangle.
\end{aligned}$$

Invoking Proposition 2.1 finishes the proof of (2.3). □

2.2 Modulation

Recall that $X^s := \dot H^{s+1} \cap \dot H^1$. Let $\boldsymbol{u}_0^* \in X^s \times H^s$, $T_+ \in \mathbb{R}$ and let $\boldsymbol{u}^*$ be the solution of (NLW) with initial data $\boldsymbol{u}^*(T_+) = \boldsymbol{u}_0^*$. Without loss of generality we will assume that $\frac{N-2}{2} < s \leq 2$. For fixed $\rho > 0$ we denote

$$c^* := \|u^*\|_{L^\infty((T_+ - \rho, T_+)\times B(0,\rho))}.$$

We can assume that $c^* > 0$ (otherwise there is no blow-up, cf. Remark 1.1). Note that because of finite speed of propagation, we can also assume that $\|\boldsymbol{u}^*(t)\|_{\mathcal{E}}$ is smaller than any fixed strictly positive constant and that $\|u^*(t)\|_{L^\infty} \leq 2c^*$ for t close to T_+.

Because of a slow decay of W, we will introduce compactly supported approximations of W_λ. Let

$$R := (c_0 \cdot c^*)^{\frac{1}{-N+2}}, \tag{2.4}$$

where $c_0 > 0$ is a small universal constant to be chosen later.

We denote

$$V(\lambda)(x) := \begin{cases} W_\lambda(x) - \zeta(\lambda) & \text{for } |x| \leq R\sqrt{\lambda}, \\ 0 & \text{for } |x| \geq R\sqrt{\lambda}, \end{cases}$$

where

$$\zeta(\lambda) := W_\lambda(R\sqrt{\lambda}) = \frac{1}{\lambda^{\frac{N-2}{2}}}\Big(1 + \frac{R^2}{N(N-2)\lambda}\Big)^{-\frac{N-2}{2}} = \Big(\lambda + \frac{R^2}{N(N-2)}\Big)^{-\frac{N-2}{2}}.$$

We will also denote

$$\boldsymbol{V}(\lambda) := (V(\lambda), 0) \in \mathcal{E}.$$

Notice that

$$\partial_\lambda V(\lambda)(x) = \begin{cases} -(\Lambda W)_{\underline{\lambda}}(x) - \zeta'(\lambda) & \text{for } |x| < R\sqrt{\lambda}, \\ 0 & \text{for } |x| > R\sqrt{\lambda}. \end{cases}$$

Lemma 2.3. *Let $s > \frac{N-2}{2}$ and $s \geq 1$. The following estimates are true with universal constants:*

$$\|V(\lambda) - W_\lambda\|_{\dot H^1} \lesssim R^{\frac{-N+2}{2}}\lambda^{\frac{N-2}{4}}, \tag{2.5}$$

$$\|V(\lambda) - W_\lambda\|_{L^\infty} \lesssim R^{-N+2}, \tag{2.6}$$

$$\|\partial_\lambda V(\lambda) + \Lambda W_{\underline{\lambda}}\|_{L^\infty(|x|<R\sqrt{\lambda})} \lesssim R^{-N}, \tag{2.7}$$

$$\|\partial_\lambda V(\lambda)\|_{L^{\frac{2N}{N+2}}} \lesssim R^{\frac{6-N}{2}}\lambda^{\frac{N-2}{4}}, \tag{2.8}$$

$$\|\partial_\lambda V(\lambda)\|_{H^{1-s}} \ll \lambda^{\frac{N-4}{2}} \qquad \textit{as } \lambda \to 0. \tag{2.9}$$

Proof. To prove (2.5), we write

$$\|V(\lambda) - W_\lambda\|_{\dot H^1}^2 = \int_{|x|\geq R\sqrt{\lambda}} |\nabla W_\lambda|^2 \,\mathrm{d}x = \int_{|x|\geq R/\sqrt{\lambda}} \|\nabla W\|^2 \,\mathrm{d}x$$
$$\lesssim \int_{R/\sqrt{\lambda}}^{+\infty} r^{-2N+2}\cdot r^{N-1}\,\mathrm{d}r \sim (R/\sqrt{\lambda})^{-N+2}.$$

We see that $\zeta(\lambda) \sim R^{-(N-2)}$ and $\zeta'(\lambda) \sim R^{-N}$ when λ is small, which proves (2.6) and (2.7).

On the support of $\partial_\lambda V(\lambda)$ there holds $|\partial_\lambda V(\lambda)(x)| \lesssim \lambda^{\frac{N-4}{2}}|x|^{-N+2}$, hence

$$\|\partial_\lambda V(\lambda)\|_{L^{\frac{2N}{N+2}}}^{\frac{2N}{N+2}} \lesssim \int_0^{R\sqrt{\lambda}} \lambda^{\frac{N-4}{2}\cdot\frac{2N}{N+2}} r^{(-N+2)\frac{2N}{N+2}} r^{N-1}\,\mathrm{d}r$$
$$= \lambda^{\frac{N^2-4N}{N+2}} \int_0^{R\sqrt{\lambda}} r^{\frac{-N^2+5N-2}{N+2}}\,\mathrm{d}r = R^{\frac{N(6-N)}{N+2}}\lambda^{\frac{N(N-2)}{2(N+2)}}.$$

This proves (2.8).

We will check (2.9) separately in each dimension. For $N = 3$ we have $|\partial_\lambda V(\lambda)(x)| \lesssim \lambda^{-\frac12}|x|^{-1}$ and $\||x|^{-1}\|_{L^2(|x|\leq R\sqrt{\lambda})} \ll 1$. For $N = 4$ we have $|\partial_\lambda V(\lambda)(x)| \lesssim |x|^{-2}$. We suppose $s > 1$, hence there exists $q \in (1,2)$ such that $L^q \subset H^{1-s}$ and it is easy to check that $\||x|^{-2}\|_{L^q(|x|\leq R\sqrt{\lambda})} \ll 1$. Finally for $N = 5$ we have $|\partial_\lambda V(\lambda)(x)| \lesssim \lambda^{\frac12}|x|^{-3}$. There exists $q \in \left(1, \frac53\right)$ such that $L^q \subset H^{1-s}$ and it is easy to check that $\||x|^{-3}\|_{L^q(|x|\leq R\sqrt{\lambda})} \ll 1$. □

Note that $\zeta(\lambda) \sim c_0c^*$, which means that the cut-off is made at a radius $r = R\sqrt{\lambda}$ such that $W_\lambda(r) \sim c_0u^*(t,r)$.

For the next lemma we will need the following version of the Implicit Function Theorem. It is obtained directly from standard proofs of the usual version, see for example [11, Section 2.2].

Lemma 2.4. *Suppose that X, Y and Z are Banach spaces, $x_0 \in X$, $y_0 \in Y$, $\rho, \eta > 0$ and $\Phi : B(x_0,\rho)\times B(y_0,\eta) \to Z$ is continuous in x and continuously differentiable in y, $\Phi(x_0,y_0) = 0$ and $\mathrm{D}_y\Phi(x_0,y_0) =: L_0$ has a bounded inverse. Suppose that*

$$\|L_0 - \mathrm{D}_y\Phi(x,y)\|_Z \leq \frac13\|L_0^{-1}\|_{\mathscr{L}(Z,Y)}^{-1} \qquad \textit{for } \|x-x_0\|_X < \rho, \|y-y_0\|_Y < \eta,$$
$$\|\Phi(x,y_0)\|_Z \leq \frac{\eta}{3}\|L_0^{-1}\|_{\mathscr{L}(Z,Y)}^{-1} \qquad \textit{for } \|x-x_0\|_X < \rho.$$

Then there exists $y \in C(B(x_0,\rho), B(y_0,\eta))$ such that for $x \in B(x_0,\rho)$, $y(x)$ is the unique solution of the equation $\Phi(x,y(x)) = 0$ in $B(y_0,\eta)$. □

Lemma 2.5. *There exists $\delta_0 > 0$ and $\lambda_0 > 0$ such that for any $0 \leq \delta \leq \delta_0$ and $t_1 < t_2$, if $\boldsymbol{u} : (t_1,t_2) \to \mathcal{E}$ is a solution of* (NLW) *satisfying for all $t \in (t_1,t_2)$:*

$$\|\boldsymbol{u}(t) - \boldsymbol{u}^*(t) - \boldsymbol{W}_{\widetilde\lambda(t)}\|_{\mathcal{E}} \leq \delta, \qquad 0 < \widetilde\lambda(t) < \lambda_0, \tag{2.10}$$

then there exists a unique function $\lambda(t) \in C^1((t_1, t_2), (0, +\infty))$ *such that*

$$\boldsymbol{g}(t) := \boldsymbol{u}(t) - \boldsymbol{u}^*(t) - \boldsymbol{V}(\lambda(t)) \tag{2.11}$$

satisfies for all $t \in (t_1, t_2)$*:*

$$\langle \underline{\mathcal{Z}_{\lambda(t)}}, g(t)\rangle = 0, \tag{2.12}$$

$$\|\boldsymbol{g}(t)\|_{\mathcal{E}} \lesssim \delta + \widetilde{\lambda}(t)^{\frac{N-2}{4}}, \tag{2.13}$$

$$|\lambda(t)/\widetilde{\lambda}(t) - 1| \lesssim \delta, \tag{2.14}$$

$$|\lambda'(t)| \lesssim \|\boldsymbol{g}(t)\|_{\mathcal{E}}. \tag{2.15}$$

Proof. We will first show that for $t_0 \in (t_1, t_2)$ fixed there exists a unique $\lambda(t_0)$ such that (2.12), (2.13) and (2.14) hold at $t = t_0$. The proof is standard, see for example [57, Proposition 1].

Denote $\boldsymbol{v}_0 := \boldsymbol{u}(t_0) - \boldsymbol{u}^*(t_0)$ and $\widetilde{l}_0 := \log(\widetilde{\lambda}(t_0))$ (it will be convenient to consider $\widetilde{\lambda}(t_0)$ and $\lambda(t_0)$ in the logarithmic scale). We define the following functional:

$$\Phi : \mathcal{E} \times \mathbb{R} \to \mathbb{R}, \qquad \Phi(\boldsymbol{v}; l) := \langle \mathrm{e}^{-l}\underline{\mathcal{Z}_{\mathrm{e}^l}}, v - V(\mathrm{e}^l)\rangle.$$

We have

$$\partial_l \Phi(\boldsymbol{v}; l) = -\langle \underline{\mathcal{Z}_{\mathrm{e}^l}}, \partial_\lambda V(\mathrm{e}^l)\rangle - \langle \mathrm{e}^{-l}\Lambda_{-1}\underline{\mathcal{Z}_{\mathrm{e}^l}}, v - V(\mathrm{e}^l)\rangle.$$

We apply Lemma 2.4 with $x_0 = \boldsymbol{V}(\widetilde{\lambda}(t_0))$ and $y_0 = \widetilde{l}_0$. It is easily checked that the assumptions hold if δ is small and $\eta = C\delta$, with a large constant C. Take $\lambda(t_0) = \mathrm{e}^{l_0}$, where l_0 is the solution of $\Phi(\boldsymbol{v}_0; l_0) = 0$ given by Lemma 2.4. Directly from the definition of Φ we obtain (2.12). The inequality $|l_0 - \widetilde{l}_0| \leq \eta = C\delta$ is equivalent to (2.14), which in turn implies

$$\|W_{\widetilde{\lambda}(t_0)} - W_{\lambda(t_0)}\|_{\dot{H}^1} \lesssim \delta. \tag{2.16}$$

From the definition of $\boldsymbol{g}$ and (2.10) we have

$$\|\boldsymbol{g}\| \leq \delta + \|W_{\widetilde{\lambda}(t_0)} - W_{\lambda(t_0)}\|_{\dot{H}^1} + \|W_{\lambda(t_0)} - V(\lambda(t_0))\|_{\dot{H}^1},$$

so (2.13) follows from (2.16) and (2.5).

For each $t_0 \in (t_1, t_2)$ we have defined $\lambda(t_0)$. It remains to show that $\lambda(t)$ is a C^1 function and that (2.15) holds. One way is to use a regularization procedure as in [57]. Here we give a different argument, which might be simpler in some cases.

Take $t_0 \in (t_1, t_2)$ and let $l_0 := \log(\lambda(t_0))$. Denote $\boldsymbol{v}(t) := \boldsymbol{u}(t) - \boldsymbol{u}^*(t)$ and define $l : (t_0 - \varepsilon, t_0 + \varepsilon) \to \mathbb{R}$ as the solution of the differential equation

$$l'(t) = -(\partial_l \Phi)^{-1}(\mathrm{D}_{\boldsymbol{v}}\Phi)\partial_t \boldsymbol{v}(t)$$

with the initial condition $l(t_0) = l_0$. Notice that $\mathrm{D}_{\boldsymbol{v}}\Phi$ is a continuous functional on $\mathcal{F}$, so we can apply it to $\partial_t \boldsymbol{v}(t)$.

Using the chain rule we get $\frac{\mathrm{d}}{\mathrm{d}t}\Phi(\boldsymbol{v}(t); l(t)) = 0$ for $t \in (t_0 - \varepsilon, t_0 + \varepsilon)$. By continuity, $|l(t) - l_0| < \eta = C\delta$ in some neighbourhood of $t = t_0$. Hence, by the uniqueness part of Lemma 2.4, we get $l(t) = \log \lambda(t)$ in some neighbourhood of $t = t_0$. In particular, $\lambda(t)$ is of class C^1 in some neighbourhood of t_0.

From (2.11) we obtain the following differential equation for the error term $\boldsymbol{g}$:

$$\partial_t \boldsymbol{g} = J \circ (\mathrm{D}E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^* + \boldsymbol{g}) - \mathrm{D}E(\boldsymbol{u}^*)) - \lambda'\partial_\lambda \boldsymbol{V}(\lambda), \tag{2.17}$$

which can also be written in the expanded form

$$\begin{cases} \partial_t g = \dot g - \lambda'\partial_\lambda V(\lambda), \\ \partial_t \dot g = \Delta g + \big(f(u^* + V(\lambda) + g) - f(u^*) - f(V(\lambda))\big) + \big(\Delta V(\lambda) + f(V(\lambda))\big). \end{cases} \tag{2.18}$$

Differentiating (2.12) and using the first equation in (2.18) we get

$$\begin{aligned} 0 = \frac{\mathrm{d}}{\mathrm{d}t}\langle \mathcal{Z}_{\underline{\lambda}}, g\rangle &= -\frac{\lambda'}{\lambda}\langle \Lambda_0 \mathcal{Z}_{\underline{\lambda}}, g\rangle + \langle \mathcal{Z}_{\underline{\lambda}}, \dot g - \lambda'\partial_\lambda V(\lambda)\rangle \\ &= \lambda'\big(\langle \mathcal{Z}, \Lambda W\rangle - \langle \mathcal{Z}_{\underline{\lambda}}, \Lambda W_{\underline{\lambda}} + \partial_\lambda V(\lambda)\rangle - \big\langle \frac{1}{\lambda}\Lambda_0\mathcal{Z}_{\underline{\lambda}}, g\big\rangle\big) + \langle \mathcal{Z}_{\underline{\lambda}}, \dot g\rangle. \end{aligned}$$

We assumed that $\langle \mathcal{Z}, \Lambda W\rangle > 0$. When $\|\boldsymbol{g}\|_\mathcal{E}$ and λ are small enough, then

$$\big|\langle \mathcal{Z}_{\underline{\lambda}}, \Lambda W_{\underline{\lambda}} + \partial_\lambda V(\lambda)\rangle + \big\langle \frac{1}{\lambda}\Lambda_0\mathcal{Z}_{\underline{\lambda}}, g\big\rangle\big| \leq \frac{1}{2}\langle \mathcal{Z}, \Lambda W\rangle$$

(we use (2.7) in order to estimate the first term). This proves (2.15). ☐

If $\boldsymbol{u}(t)$ is a solution of (NLW) satisfying (1.3), then there exists t_0 such that (2.10) holds for $t \in [t_0, T_+)$. It follows from (2.14) that, while proving Theorem 1, without loss of generality we can assume that $\lambda(t)$ is the function given by Lemma 2.5. From (2.13) we obtain that $\|\boldsymbol{g}\|_\mathcal{E} \to 0$ as $t \to T_+$, which is the only information about $\boldsymbol{g}$ used in the sequel. The precise form of the right hand side of (2.13) has no importance. We will prove that (1.4) holds on some interval $[t_0, T_+)$ with $t_0 < T_+$, with no information about the length of this interval. Each time we state something for $t \in [t_0, T_+)$ it should be understood that t_0 is sufficiently close to T_+.

In the rest of this paper $\lambda(t)$ always stands for the modulation parameter obtained in Lemma 2.5 and $\boldsymbol{g}(t)$ is the function defined by (2.11). We introduce the following notation for the joint size of the error and the interaction:

$$n(\boldsymbol{g}, \lambda) := \sqrt{\|\boldsymbol{g}\|_\mathcal{E}^2 + c^*\lambda^{\frac{N-2}{2}}}.$$

We will now analyze the stable and unstable directions of the linearized flow. The stable coefficient $a^-(t)$ and the unstable coefficient $a^+(t)$ are defined as follows:

$$a^-(t) := \langle \alpha^-_{\lambda(t)}, \boldsymbol{g}(t)\rangle, \qquad a^+(t) := \langle \alpha^+_{\lambda(t)}, \boldsymbol{g}(t)\rangle.$$

Note that $|a^-(t)| \lesssim \|\boldsymbol{g}\|_\mathcal{E}$ and $|a^+(t)| \lesssim \|\boldsymbol{g}\|_\mathcal{E}$.

Lemma 2.6. *The functions $a^-(t)$ and $a^+(t)$ satisfy*

$$\begin{aligned} &\big|\frac{\mathrm{d}}{\mathrm{d}t}a^-(t) + \frac{\nu}{\lambda(t)}a^-(t)\big| \lesssim \frac{1}{\lambda(t)}n(\boldsymbol{g}(t), \lambda(t))^2, \\ &\big|\frac{\mathrm{d}}{\mathrm{d}t}a^+(t) - \frac{\nu}{\lambda(t)}a^+(t)\big| \lesssim \frac{1}{\lambda(t)}n(\boldsymbol{g}(t), \lambda(t))^2. \end{aligned} \tag{2.19}$$

Proof. We will only prove (2.19); the other estimate can be shown analogously.

Let us rewrite equation (2.18) in the following manner:

$$\partial_t \boldsymbol{g} = J \circ \mathrm{D}^2 E(\boldsymbol{W}_\lambda)\boldsymbol{g} + \boldsymbol{h},$$

where

$$\boldsymbol{h} = \begin{pmatrix} h \\ \dot h \end{pmatrix} = \begin{pmatrix} -\lambda'\partial_\lambda V(\lambda), \\ \big(f(u^* + V(\lambda) + g) - f(u^*) - f(V(\lambda)) - f'(W_\lambda)g\big) + \big(\Delta V(\lambda) + f(V(\lambda))\big) \end{pmatrix}.$$

Using (2.2) we get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}a^-(t) + \frac{\nu}{\lambda}a^-(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\langle\alpha_\lambda^-, \boldsymbol{g}\rangle + \frac{\nu}{\lambda}\langle\alpha_\lambda^-, \boldsymbol{g}\rangle \\ &= -\frac{\lambda'}{\lambda}\langle\Lambda_{\mathcal{E}^*}\alpha_\lambda^-, \boldsymbol{g}\rangle + \langle\alpha_\lambda^-, J\circ \mathrm{D}^2E(\boldsymbol{W}_\lambda)\boldsymbol{g}\rangle + \frac{\nu}{\lambda}\langle\alpha_\lambda^-, \boldsymbol{g}\rangle + \langle\alpha_\lambda^-, \boldsymbol{h}\rangle \\ &= -\frac{\lambda'}{\lambda}\langle\Lambda_{\mathcal{E}^*}\alpha_\lambda^-, \boldsymbol{g}\rangle + \langle\alpha_\lambda^-, \boldsymbol{h}\rangle.\end{aligned}$$

The first term is negligible due to (2.15). In order to bound the second term it suffices to check the following inequalities:

$$\big|\big\langle\mathcal{Y}_{\underline{\lambda}}, \partial_\lambda V(\lambda)\big\rangle\big| \lesssim n(\boldsymbol{g},\lambda)^2,$$

$$\big|\big\langle\mathcal{Y}_\lambda, \big(\Delta V(\lambda) + f(V(\lambda))\big)\big\rangle\big| \lesssim n(\boldsymbol{g},\lambda)^2, \tag{2.20}$$

$$\big|\big\langle\mathcal{Y}_\lambda, \big(f(u^* + V(\lambda) + g) - f(u^*) - f(V(\lambda)) - f'(W_\lambda)g\big)\big\rangle\big| \lesssim n(\boldsymbol{g},\lambda)^2. \tag{2.21}$$

The first inequality follows from (2.7) and (2.1), since the region $|x| \geq R\sqrt{\lambda}$ is negligible due to exponential decay of $\mathcal{Y}$.

Notice that $|f(W_\lambda) - f(V(\lambda))| \lesssim f'(W_\lambda)| \cdot |W_\lambda - V(\lambda)| \lesssim f'(W_\lambda)c_0c^*$, where the last inequality follows from (2.5) and (2.4). Together with the fact that $\Delta(W_\lambda) + f(W_\lambda) = 0$ this implies

$$\begin{aligned}\big|\big\langle\mathcal{Y}_\lambda, \big(\Delta V(\lambda) + f(V(\lambda))\big)\big\rangle\big| &\lesssim \big|\big\langle\mathcal{Y}_\lambda, \Delta\big(W_\lambda - V(\lambda)\big)\big\rangle\big| + \big|\big\langle\mathcal{Y}_\lambda, f(W_\lambda) - f(V(\lambda))\big\rangle\big| \\ &\lesssim \big(\|\Delta Y_\lambda\|_{L^1} + \|f'(W_\lambda)\mathcal{Y}_\lambda\|_{L^1}\big)c_0c^* \lesssim c^*\lambda^{\frac{N-2}{2}},\end{aligned}$$

which proves (2.20).

We will check (2.21) in three small steps. As before, we do not have to worry about the region $|x| \geq R\sqrt{\lambda}$ thanks to the fast decay of $\mathcal{Y}$. First, we have a pointwise bound

$$|f(u^* + V(\lambda)) - f(u^*) - f(V(\lambda))| \lesssim f'(W_\lambda)\cdot c^* + f(c^*), \tag{2.22}$$

which implies

$$\big|\big\langle\mathcal{Y}_\lambda, f(u^* + V(\lambda)) - f(u^*) - f(V(\lambda))\big\rangle\big| \lesssim n(\boldsymbol{g},\lambda)^2. \tag{2.23}$$

Next, we have

$$|f(u^* + V(\lambda) + g) - f(u^* + V(\lambda)) - f'(u^* + V(\lambda))g| \lesssim |f''(u^* + V(\lambda))|\cdot|g|^2 + f(|g|), \tag{2.24}$$

which implies

$$\big|\big\langle\mathcal{Y}_\lambda, f(u^* + V(\lambda) + g) - f(u^* + V(\lambda)) - f'(u^* + V(\lambda))g\big\rangle\big| \lesssim n(\boldsymbol{g},\lambda)^2. \tag{2.25}$$

Finally, $|f'(V(\lambda) + u^*) - f'(W_\lambda)| \lesssim (|f''(W_\lambda)| + |f''(V(\lambda) + u^* - W_\lambda)|)\cdot|V(\lambda) + u^* - W_\lambda| \lesssim |f''(W_\lambda)|c^*$. Using Hölder and the fact that $\|\mathcal{Y}_\lambda \cdot f''(W_\lambda)\|_{L^{\frac{2N}{N+2}}} \lesssim \lambda^{\frac{N-2}{2}}$ this implies

$$\big|\big\langle\mathcal{Y}_\lambda, \big(f'(u^* + V(\lambda)) - f'(W_\lambda)\big)g\big\rangle\big| \ll n(\boldsymbol{g},\lambda)^2. \tag{2.26}$$

Now (2.21) follows from (2.23), (2.25) and (2.26) and the triangle inequality. □

2.3 Coercivity

By the conservation of energy, for all $t \in [t_0, T_+)$ there holds

$$E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^* + \boldsymbol{g}) = E(\boldsymbol{W}) + E(\boldsymbol{u}^*). \tag{2.27}$$

On the other hand, using the pointwise inequality

$$|F(k+l) - F(k) - f(k)l - \frac{1}{2}f'(k)l^2| \lesssim |f''(k)||l^3| + |F(l)|, \qquad \forall k, l \in \mathbb{R}$$

we deduce that

$$\begin{aligned} E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^* + \boldsymbol{g}) &= E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^*) + \langle \mathrm{D}E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^*), \boldsymbol{g}\rangle \\ &\quad + \frac{1}{2}\langle \mathrm{D}^2 E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^*)\boldsymbol{g}, \boldsymbol{g}\rangle + O(\|\boldsymbol{g}\|_{\mathcal{E}}^3). \end{aligned}$$

Using (2.27) we obtain

$$\begin{aligned} \big(E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^*) - E(\boldsymbol{W}) - E(\boldsymbol{u}^*)\big) + \langle \mathrm{D}E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^*), \boldsymbol{g}\rangle \\ + \frac{1}{2}\langle \mathrm{D}^2 E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^*)\boldsymbol{g}, \boldsymbol{g}\rangle = O(\|\boldsymbol{g}\|_{\mathcal{E}}^3). \end{aligned} \tag{2.28}$$

We start by computing the size of the first term on the left hand side.

Lemma 2.7. *For $T_+ - t$ small there holds*

$$|E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^*) - E(\boldsymbol{W}) - E(\boldsymbol{u}^*)| \lesssim c^* \lambda^{\frac{N-2}{2}}.$$

In addition, if $u^*(0) < 0$, then

$$E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^*) - E(\boldsymbol{W}) - E(\boldsymbol{u}^*) \gtrsim c^* \lambda^{\frac{N-2}{2}}. \tag{2.29}$$

Proof. Integrating by parts we obtain

$$\begin{aligned} \int \nabla V(\lambda) \cdot \nabla u^* \,\mathrm{d}x &= \int_{B(0,R\sqrt{\lambda})} \nabla(W_\lambda) \cdot \nabla u^* \,\mathrm{d}x \\ &= -\int_{B(0,R\sqrt{\lambda})} \Delta(W_\lambda) \cdot u^* \,\mathrm{d}x + \int_{S(0,R\sqrt{\lambda})} \partial_r(W_\lambda) \cdot u^* \,\mathrm{d}\sigma \\ &= \int_{B(0,R\sqrt{\lambda})} f(W_\lambda) \cdot u^* \,\mathrm{d}x + \int_{S(0,R\sqrt{\lambda})} \partial_r(W_\lambda) \cdot u^* \,\mathrm{d}\sigma. \end{aligned}$$

Developping the energy gives

$$\begin{aligned} E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^*) - E(\boldsymbol{W}) - E(\boldsymbol{u}^*) &= \int \nabla V(\lambda) \cdot \nabla u^* \,\mathrm{d}x + \frac{1}{2}\int |\nabla V(\lambda)|^2 - |\nabla(W_\lambda)|^2 \,\mathrm{d}x \\ &\quad - \int F(V(\lambda) + u^*) - F(W_\lambda) - F(u^*) \,\mathrm{d}x \\ &= \int_{S(0,R\sqrt{\lambda})} \partial_r(W_\lambda) \cdot u^* \,\mathrm{d}\sigma + \frac{1}{2}\int |\nabla V(\lambda)|^2 - |\nabla W_\lambda|^2 \,\mathrm{d}x \\ &\quad - \int F(V(\lambda) + u^*) - F(W_\lambda) - F(u^*) - f(V(\lambda)) \cdot u^* \,\mathrm{d}x \\ &\quad + \int_{B(0,R\sqrt{\lambda})} \big(f(W_\lambda) - f(V(\lambda))\big) \cdot u^* \,\mathrm{d}x. \end{aligned} \tag{2.30}$$

We will show that all the terms on the right hand side except for the first one are $\lesssim c_0 c^* \lambda^{\frac{N-2}{2}}$, where c_0 is the small constant in (2.4).

The fact that $\int |\nabla V(\lambda)|^2 - |\nabla W_\lambda|^2 \,\mathrm{d}x \lesssim c_0 c^* \lambda^{\frac{N-2}{2}} = R^{-N+2}\lambda^{\frac{N-2}{2}}$ follows directly from the proof of (2.5).

We will now show that

$$\int |F(V(\lambda)+u^*) - F(V(\lambda)) - F(u^*) - f(V(\lambda))u^*| \,\mathrm{d}x \ll \lambda^{\frac{N-2}{2}}.$$

To this end, notice first that the integrand equals 0 for $|x| \geq R\sqrt{\lambda}$. In the region $|x| \leq R\sqrt{\lambda}$ we use the pointwise estimate

$$|F(V(\lambda)+u^*) - F(V(\lambda)) - F(u^*) - f(V(\lambda))u^*| \lesssim f'(V(\lambda))|u^*|^2 + F(u^*).$$

The term $F(u^*)$ can be neglected (it is bounded in L^∞, so its contribution is at most $\lambda^{\frac{N}{2}} \ll \lambda^{\frac{N-2}{2}}$). As for the first term, it is easily checked that

$$\int_{|x|\leq R\sqrt{\lambda}} f'(W_\lambda)\,\mathrm{d}x = \lambda^{N-2}\int_{|x|\leq R/\sqrt{\lambda}} f'(W)\,\mathrm{d}x \ll \lambda^{\frac{N-2}{2}}. \tag{2.31}$$

Next, we show that if R is large enough, then

$$\int |F(W_\lambda) - F(V(\lambda))| \,\mathrm{d}x \lesssim c_0 c^* \lambda^{\frac{N-2}{2}}.$$

In the region $|x| \geq R\sqrt{\lambda}$ from (2.5) and Sobolev embedding we obtain that the contribution is at most $\lambda^{\frac{N}{2}} \ll \lambda^{\frac{N-2}{2}}$. In the region $|x| \leq R\sqrt{\lambda}$ we use the bound

$$|F(W_\lambda) - F(V(\lambda))| \lesssim \zeta(\lambda)\cdot|f(W_\lambda)| + F(\zeta(\lambda)).$$

The second term is in L^∞, so its integral is at most $O(\lambda^{\frac{N}{2}}) \ll \lambda^{\frac{N-2}{2}}$. As for the first term, it is easily seen that $\int |f(W_\lambda)|\,\mathrm{d}x \lesssim \lambda^{\frac{N-2}{2}}$, and we get the conclusion if we recall that $\zeta(\lambda) \sim c_0 c^*$.

Finally, from (2.31) and the pointwise bound $|f(V(\lambda)) - f(W_\lambda)| \lesssim |\zeta(\lambda)f'(W_\lambda)| + |f(\zeta(\lambda))|$ it follows that

$$\int_{B(0,R\sqrt{\lambda})} |f(V(\lambda)) - f(W_\lambda)|\cdot|u^*|\,\mathrm{d}x \ll \lambda^{\frac{N-2}{2}}.$$

Now consider the first term on the right hand side of (2.30). We have $\partial_r(W_\lambda)(R\sqrt{\lambda}) \sim -\lambda^{\frac{N-2}{2}}(R\sqrt{\lambda})^{-N+1}$ and $|u^*| \leq c^*$ near the origin, so we get

$$\Big|\int_{S(0,R\sqrt{\lambda})} \partial_r(W_\lambda)\cdot u^*\,\mathrm{d}\sigma\Big| \lesssim c^*\lambda^{\frac{N-2}{2}}.$$

In the case $u_0^*(0) < 0$, by continuity if in the definition of c^* we choose ρ small enough, then $u^*(t,x) \leq -\frac{1}{2}c^*$ for $(t,x) \in [t_0, T_+) \times B(0,\rho)$. In particular,

$$\int_{S(0,R\sqrt{\lambda})} \partial_r(W_\lambda)\cdot u^*\,\mathrm{d}\sigma \gtrsim c^*\lambda^{\frac{N-2}{2}},$$

where the constant in this estimate is independent of c_0. The conclusion follows from (2.30) if c_0 is chosen small enough. □

We will focus at present on the second term on the left hand side of (2.28). In Lemma 2.8 we treat the simpler case $\boldsymbol{u}_0^* \in X^2 \times H^2$ and in Lemma 2.9 we prove a weaker estimate in the case $\boldsymbol{u}^* \in X^s \times H^s$, $s > \frac{N-2}{2}$, $s \geq 1$.

Lemma 2.8. *Suppose that* $\boldsymbol{u}_0^* \in X^2 \times H^2$. *Then for* $t \in [t_0, T_+)$ *there holds*

$$|\langle \mathrm{D}E(\boldsymbol{V}(\lambda(t)) + \boldsymbol{u}^*(t)), \boldsymbol{g}(t)\rangle| \lesssim \sqrt{c_0} \cdot \sup_{t \leq \tau < T_+} n(\boldsymbol{g}(\tau), \lambda(\tau))^2,$$

where c_0 *is the small constant in* (2.4).

Proof. The proof has two steps. First we will show that

$$|\langle \mathrm{D}E(\boldsymbol{V}(\lambda(t)) + \boldsymbol{u}^*(t)) - \mathrm{D}E(\boldsymbol{u}^*(t)), \boldsymbol{g}(t)\rangle| \lesssim \sqrt{c_0} \cdot n(\boldsymbol{g}(t), \lambda(t))^2 \tag{2.32}$$

and then we will check that

$$\Big|\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathrm{D}E(\boldsymbol{u}^*(t)), \boldsymbol{g}\rangle\Big| \lesssim_R n(\boldsymbol{g}(t), \lambda(t))^2. \tag{2.33}$$

Clearly, integrating (2.33) and using (2.32), we obtain the conclusion for t_0 sufficiently close to T_+. Note that the constant in (2.33) is allowed to depend on R (because $T_+ - t_0$ can also be chosen depending on R).

In order to prove (2.32), we begin by verifying that

$$\big|\langle \mathrm{D}E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^*), \boldsymbol{g}\rangle - \langle \mathrm{D}E(\boldsymbol{V}(\lambda)), \boldsymbol{g}\rangle - \langle \mathrm{D}E(\boldsymbol{u}^*), \boldsymbol{g}\rangle\big| \ll n(\boldsymbol{g}, \lambda)^2. \tag{2.34}$$

This is equivalent to

$$\int |f(V(\lambda) + u^*) - f(V(\lambda)) - f(u^*)| \cdot |g| \,\mathrm{d}x \ll n(\boldsymbol{g}, \lambda)^2.$$

By Hölder and Sobolev inequalities, it suffices to show that

$$\|f(V(\lambda) + u^*) - f(V(\lambda)) - f(u^*)\|_{L^{\frac{2N}{N+2}}} \ll \lambda^{\frac{N-2}{4}}.$$

Using (2.22) we obtain easily that the left hand side is $\lesssim \lambda^{\frac{N-2}{2}}$.

Recall that $R^{-N+2} = c_0 c^*$, hence (2.5) gives $\|W_\lambda - V(\lambda)\|_{\dot{H}^1} \lesssim \sqrt{c_0 c^*}$. Using $\Delta W_\lambda + f(W_\lambda) = 0$ and the pointwise bound $|f(W_\lambda) - f(V(\lambda))| \lesssim f'(W_\lambda) \cdot |W_\lambda - V(\lambda)|$ one gets

$$\|\Delta V(\lambda) + f(V(\lambda))\|_{\dot{H}^{-1}} \lesssim \|\Delta(W_\lambda - V(\lambda))\|_{\dot{H}^{-1}} + \|f(W_\lambda) - f(V(\lambda))\|_{L^{\frac{2N}{N+2}}} \lesssim \sqrt{c_0 c^*},$$

hence

$$|\langle \mathrm{D}E(\boldsymbol{V}(\lambda)), \boldsymbol{g}\rangle| \lesssim \sqrt{c_0} \cdot n(\boldsymbol{g}, \lambda)^2. \tag{2.35}$$

Estimate (2.32) follows from (2.34) and (2.35). Notice that until now the assumption $\boldsymbol{u}_0^* \in X^2 \times H^2$ has not been used, thus (2.32) holds also in the case $\boldsymbol{u}_0^* \in X^s \times H^s$, $s > \frac{N-2}{2}$.

We move on to the proof of (2.33). Until the end of this proof all the constants are allowed to depend on R. From (2.17) we get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathrm{D}E(\boldsymbol{u}^*), \boldsymbol{g}\rangle &= \langle \mathrm{D}^2 E(\boldsymbol{u}^*)\partial_t \boldsymbol{u}^*, \boldsymbol{g}\rangle \\ &+ \big\langle \mathrm{D}E(\boldsymbol{u}^*), J \circ \big(\mathrm{D}E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^* + \boldsymbol{g}) - \mathrm{D}E(\boldsymbol{u}^*)\big) - \lambda' \partial_\lambda \boldsymbol{V}(\lambda)\big\rangle.\end{aligned}$$

Notice that

$$\langle \mathrm{D}^2 E(\boldsymbol{u}^*)\partial_t \boldsymbol{u}^*, \boldsymbol{g}\rangle = -\langle \mathrm{D}E(\boldsymbol{u}^*), J \circ \mathrm{D}^2 E(\boldsymbol{u}^*)\boldsymbol{g}\rangle,$$

hence it suffices to verify that

$$\big|\big\langle \mathrm{D}E(\boldsymbol{u}^*), J \circ \big(\mathrm{D}E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^* + \boldsymbol{g}) - \mathrm{D}E(\boldsymbol{u}^*) - \mathrm{D}^2 E(\boldsymbol{u}^*)\boldsymbol{g}\big) - \lambda' \partial_\lambda \boldsymbol{V}(\lambda)\big\rangle\big| \lesssim n(\boldsymbol{g}, \lambda)^2.$$

Considering separately the first and the second component, cf. (2.18), we obtain that it is sufficient to verify the following bounds:

$$|\langle \Delta u^* + f(u^*), \lambda' \partial_\lambda V(\lambda)\rangle| \lesssim n(\boldsymbol{g}, \lambda)^2, \tag{2.36}$$

$$|\langle \dot u^*, f(V(\lambda) + u^* + g) - f(V_\lambda) - f(u^*) - f'(u^*)g\rangle| \lesssim n(\boldsymbol{g}, \lambda)^2, \tag{2.37}$$

$$|\langle \dot u^*, \Delta V(\lambda) + f(V(\lambda))\rangle| \lesssim n(\boldsymbol{g}, \lambda)^2. \tag{2.38}$$

We know from Appendix A that $u^*(t)$ is bounded in X^2, hence $\Delta u^* + f(u^*)$ is bounded in $L^{\frac{2N}{N-2}}$ by the Sobolev embedding. From (2.8) and Hölder inequality it follows that

$$|\langle \Delta u^* + f(u^*), \partial_\lambda V(\lambda)\rangle| \lesssim \lambda^{\frac{N-2}{4}},$$

and (2.36) follows from (2.15).

Since $\dot u^*(t)$ is bounded in $L^{\frac{2N}{N-2}}$, in order to prove (2.37) it suffices (by Hölder) to check that

$$\|f(V(\lambda) + u^* + g) - f(V(\lambda)) - f(u^*) - f'(V(\lambda) + u^*)g\|_{L^{\frac{2N}{N+2}}} \lesssim n(\boldsymbol{g}, \lambda)^2 \tag{2.39}$$

and

$$\|\dot u^* \cdot (f'(V(\lambda) + u^*) - f'(u^*))\|_{L^{\frac{2N}{N+2}}} \lesssim \lambda^{\frac{N-2}{4}}. \tag{2.40}$$

We first prove (2.40). For $|x| \geq R\sqrt{\lambda}$ the integrand equals 0, and in the region $|x| \leq R\sqrt{\lambda}$ there holds $|f'(V(\lambda))| + |f'(u^*)| \lesssim f'(W_\lambda)$.

- For $N = 3$ $\dot u^* \in H^2 \subset L^\infty$ and $\|f'(W_\lambda)\|_{L^{\frac{6}{5}}} \lesssim \lambda^{\frac{1}{2}}$.
- For $N = 4$ $\dot u^* \in H^2 \subset L^{12}$ and $\|f'(W_\lambda)\|_{L^{\frac{3}{2}}} \lesssim \lambda^{\frac{2}{3}}$.
- For $N = 5$ $\dot u^* \in H^2 \subset L^{10}$ and $\|f'(W_\lambda)\|_{L^{\frac{5}{3}}} \lesssim \lambda$.

In all three cases (2.40) follows from Hölder inequality.

By a pointwise bound we have

$$\|f(V(\lambda) + u^*) - f(V(\lambda)) - f(u^*)\|_{L^{\frac{2N}{N+2}}} \lesssim \|u^* \cdot f'(V(\lambda))\|_{L^{\frac{2N}{N+2}}} + \|f'(u^*) \cdot V(\lambda)\|_{L^{\frac{2N}{N+2}}}.$$

It is easy to check that

$$\|f'(V(\lambda))\|_{L^{\frac{2N}{N+2}}} \leq \|f'(W_\lambda)\|_{L^{\frac{2N}{N+2}}} \lesssim \lambda^{\frac{N-2}{2}}.$$

Together with (2.8) this yields

$$\|f(V(\lambda) + u^*) - f(V(\lambda)) - f(u^*)\|_{L^{\frac{2N}{N+2}}} \lesssim \lambda^{\frac{N-2}{2}},$$

and (2.39) follows from (2.24) and the Hölder inequality.

In order to prove (2.38), we write:

$$|\langle \dot u^*, \Delta V(\lambda) + f(V(\lambda))\rangle| \leq |\langle \dot u^*, \Delta(V(\lambda) - W_\lambda)\rangle| + |\langle \dot u^*, f(V(\lambda) - f(W_\lambda)\rangle|. \tag{2.41}$$

Consider the first term of (2.41). Integrating twice by parts we find

$$\begin{aligned}\int \dot u^* \cdot \Delta(V(\lambda) - W_\lambda)\,\mathrm{d}x &= \int_{|x|\geq R\sqrt{\lambda}} \nabla \dot u^* \cdot \nabla(W_\lambda)\,\mathrm{d}x \\ &= \int_{S(0,R\sqrt{\lambda})} \dot u^* \cdot \partial_r(W_\lambda)\,\mathrm{d}\sigma - \int_{|x|\geq R\sqrt{\lambda}} \dot u^* \cdot \Delta(W_\lambda)\,\mathrm{d}x.\end{aligned}$$

As for the first term, recall that $|\partial_r(W_\lambda(R\sqrt{\lambda}))| \lesssim \lambda^{\frac{N-2}{2}}$, so it suffices to notice that by the Trace Theorem $\int |\dot{u}^*|\,\mathrm{d}\sigma \ll 1$ for $\lambda \ll 1$. In order to bound the second term, we compute

$$\|f(W_\lambda)\|_{L^{\frac{2N}{N+2}}(|x|\geq R\sqrt{\lambda})} = \|f(W)\|_{L^{\frac{2N}{N+2}}(|x|\geq R/\sqrt{\lambda})} \sim \lambda^{\frac{N+2}{4}} \ll \lambda^{\frac{N-2}{2}},$$

and use Hölder.

Consider the second term of (2.41). From (2.6) we have $|f(V(\lambda)) - f(W_\lambda)| \lesssim f'(W_\lambda)$, hence:

$$\Big|\int \dot{u}^* \cdot \big(f(V(\lambda)) - f(W_\lambda)\big)\,\mathrm{d}x\Big| \lesssim \int |\dot{u}^*| \cdot f'(W_\lambda)\,\mathrm{d}x,$$

and the required bound follows from Hölder and the fact that $\|f'(W_\lambda)\|_{L^{\frac{2N}{N+2}}} \lesssim \lambda^{\frac{N-2}{2}}$. □

Lemma 2.9. *Suppose that $\boldsymbol{u}_0^* \in X^s \times H^s$, $s > \frac{N-2}{2}$ and $s \geq 1$. There exists a decomposition*

$$\langle \mathrm{D}E(\boldsymbol{V}(\lambda(t)) + \boldsymbol{u}^*(t)), \boldsymbol{g}(t)\rangle = b_1(t) + b_2(t)$$

such that for $t \in [t_0, T_+)$ there holds:

$$|b_1'(t)| \ll \lambda(t)^{\frac{N-4}{2}}\|\boldsymbol{g}\|_{\mathcal{E}}, \tag{2.42}$$

$$|b_2(t)| \lesssim \sqrt{c_0} \cdot \sup_{t\leq\tau<T_+} n(\boldsymbol{g}(\tau), \lambda(\tau))^2. \tag{2.43}$$

Proof. We take

$$b_1(t) := \langle \mathrm{D}E(\boldsymbol{u}^*(t)), \boldsymbol{g}(t)\rangle,$$
$$b_2(t) := \langle \mathrm{D}E(\boldsymbol{V}(\lambda(t)) + \boldsymbol{u}^*(t)) - \mathrm{D}E(\boldsymbol{u}^*(t)), \boldsymbol{g}(t)\rangle.$$

Estimate (2.43) is exactly (2.32).

Repeating the computation in the proof of Lemma 2.8, we see that we need to check inequalities (2.36), (2.37) and (2.38), with " $\lesssim n(\boldsymbol{g}, \lambda)^2$ " replaced by " $\ll \lambda^{\frac{N-4}{2}}\|\boldsymbol{g}\|$ ".

We know that Δu^* is bounded in H^{s-1}, hence from (2.9) we obtain $|\langle \Delta u^*, \partial_\lambda V(\lambda)\rangle| \ll \lambda^{\frac{N-4}{2}}$. Since $\|f(u^*)\|_{L^{\frac{2N}{N-2}}}$ is bounded and $\frac{N-2}{4} > \frac{N-4}{2}$, from (2.8) we get $|\langle f(u^*), \partial_\lambda V(\lambda)\rangle| \ll \lambda^{\frac{N-4}{2}}$. Using (2.15), it follows that

$$|\langle \Delta u^* + f(u^*), \lambda'\partial_\lambda V(\lambda)\rangle| \ll \lambda^{\frac{N-4}{2}}\|\boldsymbol{g}\|.$$

The proof of (2.37) applies almost without changes, but instead of (2.40) we need to check that $\|\dot{u}^* \cdot (f'(V(\lambda) + u^*) - f'(u^*))\|_{L^{\frac{2N}{N+2}}} \ll \lambda^{\frac{N-4}{2}}$, which will follow from

$$\|\dot{u}^* \cdot f'(W_\lambda)\|_{L^{\frac{2N}{N+2}}} \ll \lambda^{\frac{N-4}{2}}. \tag{2.44}$$

We check (2.44) separately for $N = 3, 4, 5$. Recall that $\dot{u}^*$ is bounded in H^s. If $N = 3$, then $\|\dot{u}^*\|_{L^6}$ and $\|f'(W_\lambda)\|_{L^{\frac{3}{2}}}$ are bounded, hence (2.44) follows from Hölder. If $N = 4$, then (by Sobolev) there exists $q > 4$ such that $\|\dot{u}^*\|_{L^q}$ is bounded. It can be checked that for $1 < p < 2$, $\|f'(W_\lambda)\|_{L^p} \ll 1$, hence (2.44) follows. If $N = 5$, then there exists $q > 5$ such that $\|\dot{u}^*\|_{L^q}$ is bounded. It can be checked that for $\frac{5}{4} < p < 2$, $\|f'(W_\lambda)\|_{L^p} \ll \sqrt{\lambda}$, hence (2.44) follows.

In the proof of (2.38) we have only used the boundedness of $\dot{u}^*$ in H^1, hence it remains valid and gives the bound

$$|\dot{u}^*, \Delta V(\lambda) + f(V(\lambda))\rangle| \lesssim \lambda^{\frac{N-2}{2}} \ll \lambda^{\frac{N-2}{4}}.$$

□

Remark 2.10. It is not excluded that Lemma 2.8 holds under the assumption $\boldsymbol{u}_0^* \in X^s \times H^s$, $s > \frac{N-2}{2}$, but I was unable to prove it because of possible oscillations of $\lambda(t)$. Note also that Lemma 2.9 could be proved for less regular $\boldsymbol{u}_0^*$ if we had some control of $\boldsymbol{g}(t)$ in suitable (for example Strichartz) norms.

Lemma 2.8 implies that if $\boldsymbol{u}_0^* \in X^2 \times H^2$, then Lemma 2.9 holds with $b_1(t) = 0$.

For $t_0 \leq t < T_+$ we define

$$\varphi(t) := C_{\mathrm{I}} c^* \lambda(t)^{\frac{N-2}{2}} - b_1(t) + 2\big(a^-(t)^2 + a^+(t)^2\big) \tag{2.45}$$

(C_{I} is a constant to be chosen shortly). From (2.28) we have

$$\begin{aligned}\varphi(t) := {} & C_{\mathrm{I}} c^* \lambda(t)^{\frac{N-2}{2}} + \big(E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^*) - E(\boldsymbol{W}) - E(\boldsymbol{u}^*)\big) \\ & + \frac{1}{2}\langle \mathrm{D}^2 E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^*)\boldsymbol{g}, \boldsymbol{g}\rangle + 2\big(a^-(t)^2 + a^+(t)^2\big) + b_2(t) + O(\|\boldsymbol{g}\|_{\mathcal{E}}^3).\end{aligned} \tag{2.46}$$

We will consider the maximal function:

$$\varphi_{\mathrm{M}}(t) := \sup_{t \leq \tau < T_+} \varphi(\tau).$$

Note that $\varphi_{\mathrm{M}} : [t_0, T_+) \to \mathbb{R}$ is decreasing, $\lim_{t \to T_+} \varphi_{\mathrm{M}}(t) = 0$ and $0 \geq \varphi_{\mathrm{M}}'(t) \geq \min(0, \varphi'(t))$ almost everywhere.

Corollary 2.11. *Let $s > \frac{N-2}{2}$ and $s \geq 1$. For $t_0 \leq t < T_+$ there holds*

$$\varphi_{\mathrm{M}}(t) \sim \sup_{t \leq \tau < T_+} n(\boldsymbol{g}(\tau), \lambda(\tau))^2.$$

Proof. Lemma 2.7 and (2.28) yield $|\langle \mathrm{D}E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^*), \boldsymbol{g}\rangle| \lesssim n(\boldsymbol{g}, \lambda)^2$, hence from Lemma 2.9 we have

$$|b_1(t)| \lesssim \sup_{t \leq \tau < T_+} n(\boldsymbol{g}(t), \lambda(t))^2. \tag{2.47}$$

Let $t \in [t_0, T_+)$ and let $t_1 \in [t, T_+)$ be such that $\varphi_{\mathrm{M}}(t) = \varphi(t_1)$ (such t_1 exists by the definition of φ_{M}). Using (2.47) we obtain

$$\varphi_{\mathrm{M}}(t) = \varphi(t_1) \lesssim \sup_{t_1 \leq \tau < T_+} n(\boldsymbol{g}(\tau), \lambda(\tau))^2 \leq \sup_{t \leq \tau < T_+} n(\boldsymbol{g}(\tau), \lambda(\tau))^2.$$

Now let $t_2 \in [t, T_+)$ be such that $\sup_{t \leq \tau < T_+} n(\boldsymbol{g}(\tau), \lambda(\tau))^2 = n(\boldsymbol{g}(t_2), \lambda(t_2))^2$. From Lemma 2.2 and the fact that $\|\boldsymbol{V}(\lambda) + \boldsymbol{u}^* - \boldsymbol{W}_\lambda\|_{\mathcal{E}}$ is small we obtain

$$\frac{1}{2}\langle \mathrm{D}^2 E(\boldsymbol{V}(\lambda(t_2)) + \boldsymbol{u}^*(t_2))\boldsymbol{g}(t_2), \boldsymbol{g}(t_2)\rangle + 2\big(a^-(t_2)^2 + a^+(t_2)^2\big) \gtrsim \|\boldsymbol{g}(t_2)\|_{\mathcal{E}}^2. \tag{2.48}$$

From Lemma 2.7, if we choose C_{I} large enough, then $C_{\mathrm{I}} c^* \lambda^{\frac{N-2}{2}} + E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^*) - E(\boldsymbol{W}) - E(\boldsymbol{u}^*) \gtrsim c^* \lambda^{\frac{N-2}{2}}$, hence (2.46) and (2.48) yield

$$\varphi(t_2) - b_2(t_2) \gtrsim n(\boldsymbol{g}(t_2), \lambda(t_2))^2.$$

From Lemma 2.9 we have $|b_2(t_2)| \leq \sqrt{c_0} \cdot \sup_{t_2 \leq \tau < T_+} n(\boldsymbol{g}(\tau), \lambda(\tau))^2 = \sqrt{c_0} \cdot n(\boldsymbol{g}(t_2), \lambda(t_2))^2$, hence we obtain

$$\varphi_{\mathrm{M}}(t) \geq \varphi(t_2) \gtrsim n(\boldsymbol{g}(t_2), \lambda(t_2))^2 = \sup_{t \leq \tau < T_+} n(\boldsymbol{g}(\tau), \lambda(\tau))^2,$$

provided that c_0 is small enough. □

2.4 Differential inequalities and conclusion

Lemma 2.12. *There exists a constant C_a such that for $T_+ - t$ small enough there holds*

$$|a^+(t)| \leq C_a \cdot \sup_{t \leq \tau < T_+} n(\boldsymbol{g}(\tau), \lambda(\tau))^2, \qquad |a^-(t)| \leq C_a \cdot \sup_{t \leq \tau < T_+} n(\boldsymbol{g}(\tau), \lambda(\tau))^2.$$

Proof. It follows from (2.19) that there exists $C_1 > 0$ such that

$$|a^+(t)| \geq C_1 \cdot n(\boldsymbol{g}(t), \lambda(t))^2 \quad \Rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t}|a^+(t)| \geq \frac{\nu}{2\lambda(t)}|a^+(t)|. \tag{2.49}$$

Suppose that

$$|a^+(t)| \geq 2C_1 \cdot \sup_{t \leq \tau < T_+} n(\boldsymbol{g}(\tau), \lambda(\tau))^2$$

and suppose that $t_1 \in [t, T_+)$ is the smallest time such that

$$|a^+(t_1)| \leq C_1 \cdot \sup_{t_1 \leq \tau < T_+} n(\boldsymbol{g}(\tau), \lambda(\tau))^2.$$

Clearly $t_1 > t$. The function on the right hand side is decreasing with respect to t_1, hence $\frac{\mathrm{d}}{\mathrm{d}t}|a^+(t)|_{t=t_1} \leq 0$. This contradicts (2.49), hence for all $t' \in [t, T_+)$ we have

$$|a^+(t')| \geq C_1 \cdot n(\boldsymbol{g}(t'), \lambda(t'))^2. \tag{2.50}$$

Observe that

$$\int_t^{T_+} \frac{1}{\lambda(\tau)}\,\mathrm{d}\tau \gtrsim \int_t^{T_+} \frac{|\lambda'(\tau)|}{\lambda(\tau)}\,\mathrm{d}\tau = +\infty. \tag{2.51}$$

From (2.50), (2.49) and (2.51) we obtain $|a^+(t)| \to +\infty$ as $t \to T_+$, a contradiction.

We will now consider $a^-(t)$, which is less straightforward. It follows from (2.19) that there exists $C_2 > 0$ such that

$$|a^-(t)| \geq C_2 \cdot n(\boldsymbol{g}(t), \lambda(t))^2 \quad \Rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t}|a^-(t)| \leq -\frac{\nu}{2\lambda(t)}|a^-(t)|. \tag{2.52}$$

From Corollary 2.11 we obtain existence of a constant $C_3 > 0$ such that

$$|a^-(t)| \geq C_3 \cdot \varphi_{\mathrm{M}}(t) \quad \Rightarrow \quad |a^-(t)| \geq C_2 \cdot n(\boldsymbol{g}(t), \lambda(t))^2 \tag{2.53}$$

and a constant $C_4 > 0$ such that

$$|a^-(t)| \geq C_4 \cdot \sup_{t \leq \tau < T_+} n(\boldsymbol{g}(t), \lambda(t))^2 \quad \Rightarrow \quad |a^-(t)| \geq 2C_3 \cdot \varphi_{\mathrm{M}}(t).$$

Suppose that $t \in [t_0, T_+)$ is such that

$$|a^-(t)| \geq C_4 \cdot \sup_{t \leq \tau < T_+} n(\boldsymbol{g}(\tau), \lambda(\tau))^2 \tag{2.54}$$

and let $t_1 \in [t_0, t]$ be the smallest time such that for $t' \in [t_1, t]$ there holds

$$|a^-(t')| \geq C_3 \cdot \varphi_{\mathrm{M}}(t'). \tag{2.55}$$

Of course $t_1 < t$. Suppose that $t_1 > t_0$. This implies

$$-\frac{C_2\nu}{\lambda(t_1)} n(\boldsymbol{g}(t_1), \lambda(t_1))^2 \geq -\frac{\nu}{2\lambda(t_1)}|a^-(t_1)| \geq \frac{\mathrm{d}}{\mathrm{d}t}|a^-(t)|_{t=t_1} \geq C_3 \cdot \varphi_{\mathrm{M}}'(t_1)$$

(we use respectively (2.53), (2.52) and the definition of t_1).

However, $|\varphi_{\mathrm{M}}'(t_1)| \leq |\varphi'(t_1)| \ll \frac{1}{\lambda(t_1)} n(\boldsymbol{g}(t_1), \lambda(t_1))^2$, as is easily seen from (2.45). The contradiction shows that $t_1 = t_0$, hence (2.55) holds for $t' \in [t_0, t]$. This means that if there exist times t arbitrarily close to T_+ such that (2.54) holds, then (2.55) is true for $t' \in [t_0, T_+)$. From (2.52) and (2.53) we deduce that for $t \in [t_0, T_+)$ there holds

$$|a^-(t)| \leq |a^-(t_0)| \cdot \exp\Big(- \int_{t_0}^t \frac{\nu \, \mathrm{d}t}{2\lambda(t)} \Big).$$

By (2.53) and (2.15), this implies

$$|\lambda'(t)| \lesssim \exp\Big(- \int_{t_0}^t \frac{\nu \, \mathrm{d}t}{4\lambda(t)} \Big).$$

Dividing both sides by $\lambda(t)$ and integrating we get a contradiction.

We have proved the lemma with $C_a := \max(2C_1, C_4)$. □

By modifying t_0 we can assume that Lemma 2.12 holds for $t \in [t_0, T_+)$.

Proof of Theorem 1. We define

$$\widetilde{\varphi}(t) := C_{\mathrm{I}} c^* \lambda(t)^{\frac{N-2}{2}} - b_1(t), \qquad \widetilde{\varphi}_{\mathrm{M}}(t) := \sup_{t \leq \tau < T_+} \widetilde{\varphi}(\tau).$$

From Lemma 2.12 and Corrolary 2.11, it is clear that

$$\widetilde{\varphi}_{\mathrm{M}}(t) \sim \sup_{t \leq \tau < T_+} n(\boldsymbol{g}(\tau), \lambda(\tau))^2. \tag{2.56}$$

We will consider first the case $N \in \{4, 5\}$. Using (2.42) and (2.56) we obtain the following differential inequality for $t \in [t_0, T_+)$:

$$|\widetilde{\varphi}_{\mathrm{M}}'(t)| \leq |\widetilde{\varphi}'(t)| \lesssim c^* \lambda(t)^{\frac{N-4}{2}} \|\boldsymbol{g}(t)\|_{\mathcal{E}} \lesssim (c^*)^{\frac{2}{N-2}} \widetilde{\varphi}_{\mathrm{M}}(t)^{\frac{3N-10}{2(N-2)}}. \tag{2.57}$$

Integrating this inequality we find

$$\widetilde{\varphi}_{\mathrm{M}}(t) \lesssim (c^*)^{\frac{4}{6-N}} (T_+ - t)^{\frac{2(N-2)}{6-N}}.$$

To finish the proof, recall that $c^* \lambda(t)^{\frac{N-2}{2}} \lesssim \widetilde{\varphi}_{\mathrm{M}}(t)$ by Corollary 2.11.

Consider now the case $N = 3$. The problem is that $N - 4 < 0$, hence we cannot write $(c^* \lambda(t))^{\frac{N-4}{2}} \lesssim \widetilde{\varphi}_{\mathrm{M}}^{\frac{N-4}{2(N-2)}}$, as we did in the previous proof. Instead, we just have

$$|\widetilde{\varphi}_{\mathrm{M}}'(t)| \lesssim c^* \lambda(t)^{-\frac{1}{2}} \cdot \sqrt{\widetilde{\varphi}_{\mathrm{M}}(t)}.$$

Integrating between t and T_+ we obtain

$$\sqrt[4]{\lambda(t)} \lesssim \sqrt{c^*} \int_t^{T_+} \frac{\mathrm{d}\tau}{\sqrt{\lambda(\tau)}}.$$

This is again a differential inequality. It yields (1.6). □

Remark 2.13. In the case $N = 3$ and $\boldsymbol{u}^* \in X^2 \times H^2$, we can prove (1.4) for continuous time, not only for a sequence. Indeed, in this case one can take $b_1(t) = 0$ (see Remark 2.10), hence $\widetilde{\varphi}_{\mathrm{M}}(t) = C_{\mathrm{I}} c^* \sqrt{\lambda(t)}$. If $t \in [t_0, T_+)$ is such that $\lambda(t) < \sup_{t \leq \tau < T_+} \lambda(\tau)$, then obviously $\widetilde{\varphi}_{\mathrm{M}}'(t) = 0$. If $\lambda(t) = \sup_{t \leq \tau < T_+} \lambda(\tau)$, then $c^* \sqrt{\lambda(t)} \sim \widetilde{\varphi}_{\mathrm{M}}(t)$, hence the proof of (2.57) applies. The end of the proof is the same as in the case $N \in \{4, 5\}$.

Proof of Theorem 2. Let $t \in [t_0, T_+)$ be such that $n(\boldsymbol{g}(t), \lambda(t)) = \sup_{t \leq \tau < T_+} n(\boldsymbol{g}(\tau), \lambda(\tau))$. From (2.29) and Lemma 2.2 we get

$$\big(E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^*) - E(\boldsymbol{W}) - E(\boldsymbol{u}^*)\big) + \frac{1}{2}\langle \mathrm{D}^2 E(\boldsymbol{V}(\lambda) + \boldsymbol{u}^*)\boldsymbol{g}, \boldsymbol{g}\rangle + 2((a^-)^2 + (a^+)^2) \gtrsim n(\boldsymbol{g}, \lambda)^2.$$

But due to Lemma 2.12, the last term on the right hand side can be omitted. This is in contradiction with (2.28) and Lemma 2.8. □

A Cauchy theory in higher regularity

In this section we prove some facts about propagation of regularity for (NLW), which are applied to $\boldsymbol{u}^*(t)$ in the main text. As in [40, Appendix B], the proofs rely on the classical *energy estimates*:

Proposition. *Let $s \geq 0$, $t_0 \in [T_1, T_2]$, $g \in L^1(I, H^s)$ and $\boldsymbol{u}_0 \in X^s \times H^s$. Then the solution of the linear wave equation $(\partial_{tt} - \Delta)u = g$ with initial data $\boldsymbol{u}(t_0) = \boldsymbol{u}_0$ satisfies*

$$\|\boldsymbol{u}(t)\|_{X^s \times H^s} \leq \|\boldsymbol{u}_0\|_{X^s \times H^s} + \Big|\int_{t_0}^t \|g(\tau)\|_{H^s}\,\mathrm{d}\tau\Big|, \qquad \forall t \in [T_1, T_2].$$

□

Proposition A.1. *Let $N \in \{3, 4\}$, $s > \frac{N-2}{2}$ and $\boldsymbol{u}_0 \in X^s \times H^s$. There exist $t_1 < t_0 < t_2$ such that the solution $\boldsymbol{u}(t)$ of* (NLW) *satisfies*

$$\boldsymbol{u} \in C([t_1, t_2], X^s \times H^s).$$

Proof. This is a standard application of the energy estimates and the Fixed Point Theorem, using the fact that $f(u)$ is a monomial and $X^s \hookrightarrow L^\infty$. We skip the details. □

In the rest of this section we consider (NLW) in dimension $N = 5$. In this case the nonlinearity $f(u) = |u|^{\frac{4}{3}}u$ is not smooth. We will use the following regularization:

$$f_n(u) := \big(1 - \chi(nu)\big)f(u), \qquad n \in \{1, 2, 3, \ldots\},$$

where

$$\chi \in C^\infty, \quad \chi(-u) = \chi(u), \quad \chi(u) = 1 \text{ for } u \in [-1, 1], \quad \operatorname{supp}\chi \subset [-2, 2].$$

In the proof of the next result we will use the Fractional Leibniz Rule and the Fractional Chain Rule in the form given in [12, Propositions 3.1, 3.3]:

Proposition A.2.

- *If $\Psi \in C^1$, $0 < \alpha < 1$ and $1 < p, p_1, p_2$ are such that $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$, then*

$$\||\nabla|^\alpha \Psi(u)\|_{L^p} \lesssim \|\Psi'(u)\|_{L^{p_1}} \cdot \||\nabla|^\alpha u\|_{L^{p_2}}.$$

- *If $0 < \alpha < 1$ and $1 < p, p_1, p_2, \widetilde{p}_1, \widetilde{p}_2$ are such that $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{\widetilde{p}_1} + \frac{1}{\widetilde{p}_2}$, then*

$$\||\nabla|^\alpha(uv)\|_{L^p} \lesssim \||\nabla|^\alpha u\|_{L^{p_1}} \cdot \|v\|_{L^{p_2}} + \|u\|_{L^{\widetilde{p}_1}} \cdot \||\nabla|^\alpha v\|_{L^{\widetilde{p}_2}}.$$

□

Remark A.3. In [12], the Leibniz Rule and the Chain Rule are proved in the case of one space dimension, and necessary changes in order to carry out a proof in arbitrary dimension are indicated. In the present paper we use this result in dimension 5, but only for radial functions, and it can be verified that the Leibniz Rule and the Chain Rule for radial functions is a consequence of the one-dimensional result.

Lemma A.4. *Let $N = 5$ and $1 \leq s \leq 2$. The following estimates hold (with constants which may depend on s):*

$$\|f(u) - f_n(u)\|_{H^1} \leq c_n\big(1 + f(\|u\|_{X^1})\big), \qquad \text{with } c_n \to 0 \text{ as } n \to +\infty, \tag{A.1}$$
$$\|f(u) - f(v)\|_{H^1} \lesssim \|u - v\|_{X^1} \cdot \big(f'(\|u\|_{X^1}) + f'(\|v\|_{X^1})\big), \tag{A.2}$$
$$\|f_n(u) - f_n(v)\|_{H^1} \lesssim \|u - v\|_{X^1} \cdot \big(f'(\|u\|_{X^1}) + f'(\|v\|_{X^1})\big), \tag{A.3}$$
$$\|f(u)\|_{H^s} \lesssim f(\|u\|_{X^s}), \tag{A.4}$$
$$\|f_n(u)\|_{H^s} \lesssim f(\|u\|_{X^s}), \tag{A.5}$$
$$\|f_n(u) - f_n(v)\|_{H^s} \leq C_n\|u - v\|_{X^s} \cdot \big(1 + f'(\|u\|_{X^s}) + f'(\|v\|_{X^s})\big), \qquad C_n > 0, \tag{A.6}$$

where the sign $\lesssim$ means that the constant is independent of n.

Proof. A simple computation shows that

$$|f_n(u)| \leq |f(u)|, \quad |f_n'(u)| \lesssim |f'(u)|, \quad |f_n''(u)| \lesssim |f''(u)|,$$
$$f_n \to f \quad \text{in } C^2(\mathbb{R}), \tag{A.7}$$
$$|f_n'''(u)| \lesssim n^{\frac{2}{3}}. \tag{A.8}$$

We have

$$\|\nabla(f(u) - f_n(u))\|_{L^2} = \|(f'(u) - f_n'(u))\nabla u\|_{L^2} \leq \|f' - f_n'\|_{L^\infty} \cdot \|u\|_{H^1},$$

which is acceptable due to (A.7).

In order to bound $\|f(u) - f_n(u)\|_{L^2}$, we interpolate between $\|f - f_n\|_{L^\infty}$ and

$$\|f(u) - f_n(u)\|_{L^{\frac{10}{7}}} \lesssim f(\|u\|_{L^{\frac{10}{3}}}) \lesssim f(\|u\|_{H^1}).$$

This proves (A.1).

Estimate (A.2) is a part of [40, Lemma B.3] and the proof of (A.3) is analogous.

From the Sobolev inequality we get $\|f_n(u)\|_{L^2} \leq \|f(u)\|_{L^2} \leq f(\|u\|_{L^{\frac{14}{3}}}) \lesssim f(\|u\|_{X^s})$, hence in order to prove (A.4) and (A.5) it suffices to check that

$$\||\nabla|^s(f(u))\|_{L^2} \lesssim f(\|u\|_{X^s}), \qquad \||\nabla|^s(f_n(u))\|_{L^2} \lesssim f(\|u\|_{X^s}).$$

For $s \in \{1, 2\}$ this is an easy algebraic computation which we will skip. For $1 < s < 2$ we use Proposition A.2:

$$\begin{aligned}\||\nabla|^s(f(u))\|_{L^2} &= \||\nabla|^{s-1}\nabla(f(u))\|_{L^2} = \||\nabla|^{s-1}(f'(u)\nabla u)\|_{L^2} \\ &\lesssim \||\nabla|^{s-1}\nabla u\|_{L^{\frac{10}{3}}} \cdot \|f'(u)\|_{L^5} + \||\nabla|^{s-1}(f'(u))\|_{L^5} \cdot \|\nabla u\|_{L^{\frac{10}{3}}} \\ &\lesssim \||\nabla|^{s-1}\nabla u\|_{H^1} \cdot f'(\|u\|_{L^{\frac{20}{3}}}) + \|f''(u)\|_{L^{10}} \cdot \||\nabla|^{s-1}u\|_{L^{10}} \cdot \|\nabla u\|_{H^1} \\ &\lesssim f(\|u\|_{X^s}).\end{aligned} \tag{A.9}$$

The second inequality in (A.9) is proved analogously.

In order to prove (A.6) it suffices to check that

$$\||\nabla|^s(f_n(u) - f_n(v))\|_{L^2} \le C_n \|u - v\|_{X^s} \cdot \big(1 + f'(\|u\|_{X^s}) + f'(\|v\|_{X^s})\big)$$

(the estimate of $\|f_n(u) - f_n(v)\|_{L^2}$ is a part of (A.3)). We write

$$f_n(u) - f_n(v) = -(v-u)\int_0^1 f_n'((1-t)u + tv)\,\mathrm{d}t,$$

hence by the triangle inequality

$$\||\nabla|^s(f_n(u) - f_n(v))\|_{L^2} \le \int_0^1 \big\||\nabla|^s\big((u-v)f_n'((1-t)u+tv)\big)\big\|_{L^2}\,\mathrm{d}t.$$

We will estimate the integrand for fixed $t \in [0,1]$. We have

$$\begin{aligned}\||\nabla|^s\big((u-v)f_n'((1-t)u+tv)\big)\big\|_{L^2} &= \||\nabla|^{s-1}\nabla\big((u-v)f_n'((1-t)u+tv)\big)\big\|_{L^2} \\ &= \big\||\nabla|^{s-1}\big(\nabla(u-v)\cdot f_n'((1-t)u+tv)\big)\big\|_{L^2} \\ &+ \big\||\nabla|^{s-1}\big((u-v)\cdot((1-t)\nabla u + t\nabla v)\cdot f_n''((1-t)u+tv)\big)\big\|_{L^2}.\end{aligned}$$

The first term is estimated exactly as in (A.9), so we will only consider the second one. From the Leibniz Rule we obtain

$$\begin{aligned}&\big\||\nabla|^{s-1}\big((u-v)\cdot((1-t)\nabla u + t\nabla v)\cdot f_n''((1-t)u+tv)\big)\big\|_{L^2} \\ \lesssim &\||\nabla|^{s-1}(u-v)\|_{L^{10}} \cdot \|(1-t)u + tv\|_{L^{\frac{10}{3}}} \cdot \|f_n''((1-t)u+tv)\|_{L^{10}} \\ +&\|u-v\|_{L^{10}} \cdot \||\nabla|^{s-1}((1-t)\nabla u + t\nabla v)\|_{L^{\frac{10}{3}}} \cdot \|f_n''((1-t)u+tv)\|_{L^{10}} \\ +&\|u-v\|_{L^{p_1}} \cdot \|(1-t)\nabla u + t\nabla v\|_{L^{p_2}} \cdot \||\nabla|^{s-1} f_n''((1-t)u+tv)\|_{L^{p_3}},\end{aligned}$$

where the exponents $p_1,\ p_2, p_3 \in (1, +\infty)$ are chosen such that $p_1 > 10$, $p_2 > \frac{10}{3}$, $p_3 < 10$, $X^s \subset L^{p_1} \cap W^{1,p_2}$ and $\frac{1}{2} = \frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3}$. Estimating the first two lines is straightforward and for the last line we use the Chain Rule together with (A.8). □

Proposition A.5. *Let $N = 5$, $1 \le s \le 2$ and $\boldsymbol{u}_0 \in X^s \times H^s$. There exist $t_1 < t_0 < t_2$ such that the solution $\boldsymbol{u}(t)$ of* (NLW) *satisfies*

$$\boldsymbol{u} \in C([t_1, t_2], X^s \times H^s).$$

Proof. Using (A.4) for $s = 1$ and (A.2) one obtains by a standard procedure that there exists a unique maximal solution

$$\boldsymbol{u} \in C([T_1, T_2], X^1 \times H^1), \qquad T_1 < t_0 < T_2$$

and

$$T_1 > -\infty \ \Rightarrow\ \lim_{t\to T_1}\|\boldsymbol{u}_n\|_{X^1\times H^1} = +\infty, \qquad T_2 < +\infty \ \Rightarrow\ \lim_{t\to T_2}\|\boldsymbol{u}_n\|_{X^1\times H^1} = +\infty,$$

see [40, Proposition B.2] for details.

Consider the regularized problem for $n \in \{1, 2, 3, \ldots\}$:

$$\begin{cases}(\partial_{tt} - \Delta)u_n = f_n(u_n), \\ (u_n(t_0), \partial_t u_n(t_0)) = \boldsymbol{u}_0.\end{cases}$$

Using (A.5) and (A.6) one can show that there exists a unique maximal solution

$$\boldsymbol{u}_n \in C([T_{1,n}, T_{2,n}], X^s \times H^s), \qquad T_{1,n} < t_0 < T_{2,n}$$

and

$$T_{1,n} > -\infty \ \Rightarrow \lim_{t\to T_{1,n}} \|\boldsymbol{u}_n\|_{X^s\times H^s} = +\infty, \qquad T_{2,n} < +\infty \ \Rightarrow \lim_{t\to T_{2,n}} \|\boldsymbol{u}_n\|_{X^s\times H^s} = +\infty.$$

From (A.5) and the energy estimate we have

$$\|\boldsymbol{u}_n(t)\|_{X^s\times H^s} \lesssim \|\boldsymbol{u}_0\|_{X^s\times H^s} + \Big|\int_{t_0}^{t} f(\|\boldsymbol{u}(\tau)\|_{X^s\times H^s})\,\mathrm{d}\tau\Big|,$$

with a constant independent of n. This implies that there exist $\widetilde{T}_1 < t_0$, $\widetilde{T}_2 > t_0$ and a constant C_1 independent of n such that

$$\|\boldsymbol{u}_n(t)\|_{X^s\times H^s} \leq C_1 \qquad \forall n,\ \forall t \in [\widetilde{T}_1, \widetilde{T}_2] \tag{A.10}$$

(in particular $\widetilde{T}_1 \geq \sup_n T_{1,n}$ and $\widetilde{T}_2 \leq \inf_n T_{2,n}$).

Now we need to verify that

$$\lim_{n\to+\infty} \|\boldsymbol{u}_n(t) - \boldsymbol{u}(t)\|_{X^1\times H^1} = 0 \qquad \forall t \in [\widetilde{T}_1, \widetilde{T}_2]. \tag{A.11}$$

To this end, we notice that $\boldsymbol{u}_n - \boldsymbol{u}$ solves the Cauchy problem:

$$\begin{cases} (\partial_{tt} - \Delta)(u_n - u) = f_n(u_n) - f(u), \\ (u_n(t_0), \partial_t u_n(t_0)) = 0. \end{cases}$$

Since $\|\boldsymbol{u}(t)\|_{X^1\times H^1}$ is bounded and $\|\boldsymbol{u}_n(t)\|_{X^1\times H^1}$ are uniformly bounded for $t \in [\widetilde{T}_1, \widetilde{T}_2]$, (A.1) and (A.3) imply that for $t \in [\widetilde{T}_1, \widetilde{T}_2]$ there holds

$$\|f_n(u_n(t)) - f(u(t))\|_{H^1} \leq \|f_n(u_n(t)) - f_n(u(t))\|_{H^1} + \|f_n(u(t)) - f(u(t))\|_{H^1} \lesssim \|u_n(t) - u(t)\|_{X^1} + c_n,$$

which yields (A.11) by the energy estimate and the Gronwall inequality.

From (A.10) and (A.11) we deduce

$$\|\boldsymbol{u}(t)\|_{X^s\times H^s} \leq C_1, \qquad \forall t \in [\widetilde{T}_1, \widetilde{T}_2].$$

The function $\boldsymbol{u} : [\widetilde{T}_1, \widetilde{T}_2] \to X^s \times H^s$ is weakly measurable (since it is measurable as a function to $X^1 \times H^1$), hence it is measurable and $\boldsymbol{u} \in L^\infty([\widetilde{T}_1, \widetilde{T}_2], X^s \times H^s)$. Using once again the energy estimate together with (A.4) it is easy to see that in fact $\boldsymbol{u} \in C([\widetilde{T}_1, \widetilde{T}_2], X^s \times H^s)$.

□

Chapter 4

Nonexistence of radial two-bubbles with opposite signs for the energy-critical wave equation

Abstract

We consider the focusing energy-critical wave equation in space dimension $N \geq 3$ for radial data. We study two-bubble solutions, that is solutions which behave as a superposition of two decoupled radial ground states (called bubbles) asymptotically for large positive times. We prove that in this case these two bubbles must have the same sign. The main tool is a sharp coercivity property of the energy functional near the family of ground states.

1 Introduction

1.1 Setting of the problem and the main result

Let $N \geq 3$ be the dimension of the space. For $\boldsymbol{u}_0 = (u_0, \dot u_0) \in \mathcal{E} := \dot H^1(\mathbb{R}^N) \times L^2(\mathbb{R}^N)$, define the *energy functional*

$$E(\boldsymbol{u}_0) = \int \frac 12 |\dot u_0|^2 + \frac 12 |\nabla u_0|^2 - F(u_0)\,\mathrm{d}x,$$

where $F(u_0) := \frac{N-2}{2N}|u_0|^{\frac{2N}{N-2}}$. Note that $E(\boldsymbol{u}_0)$ is well-defined due to the Sobolev Embedding Theorem. The differential of E is $\mathrm{D}E(\boldsymbol{u}_0) = (-\Delta u_0 - f(u_0), \dot u_0)$, where $f(u_0) = |u_0|^{\frac{4}{N-2}}u_0$.

We consider the Cauchy problem for the energy critical wave equation:

$$\begin{cases} \partial_t \boldsymbol{u}(t) = J\circ \mathrm{D}E(\boldsymbol{u}(t)), \\ \boldsymbol{u}(t_0) = \boldsymbol{u}_0 \in \mathcal{E}. \end{cases} \tag{NLW}$$

Here, $J := \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0\end{pmatrix}$ is the natural symplectic structure. This equation is often written in the form

$$\partial_{tt} u = \Delta u + f(u).$$

Equation (NLW) is locally well-posed in the space $\mathcal{E}$, see for example Ginibre, Soffer and Velo [32], Shatah and Struwe [84] (the defocusing case), as well as a complete review of the Cauchy theory in Kenig and Merle [47] (for $N \in \{3,4,5\}$), as well as Bulut, Czubak, Li, Pavlović and Zhang [8] (for $N \geq 6$). In particular, for any initial data $\boldsymbol{u}_0 \in \mathcal{E}$ there exists a maximal time of existence (T_-, T_+), $-\infty \leq T_- < t_0 < T_+ \leq +\infty$, and a unique solution $\boldsymbol{u} \in C((T_-,T_+);\mathcal{E})$. In addition, the energy E is a conservation law. In this paper we always assume that the initial data is radially symmetric. This symmetry is preserved by the flow.

For functions $v \in \dot H^1$, $\dot v \in L^2$, $\boldsymbol{v} = (v, \dot v) \in \mathcal{E}$ and $\lambda > 0$, we denote

$$v_\lambda(x) := \frac{1}{\lambda^{(N-2)/2}} v\big(\frac x\lambda\big), \qquad \dot v_{\underline\lambda}(x) := \frac{1}{\lambda^{N/2}}\dot v\big(\frac x\lambda\big), \qquad \boldsymbol{v}_\lambda(x) := \big(v_\lambda, \dot v_{\underline\lambda}\big).$$

A change of variables shows that

$$E\big((\boldsymbol{u}_0)_\lambda\big) = E(\boldsymbol{u}_0).$$

Equation (NLW) is invariant under the same scaling: if $\boldsymbol{u}(t) = (u(t), \dot u(t))$ is a solution of (NLW) and $\lambda > 0$, then $t \mapsto \boldsymbol{u}\big((t-t_0)/\lambda\big)_\lambda$ is also a solution with initial data $(\boldsymbol{u}_0)_\lambda$ at time $t = 0$. This is why equation (NLW) is called *energy-critical.*

A fundamental object in the study of (NLW) is the family of stationary solutions $\boldsymbol{u}(t) \equiv \pm \boldsymbol{W}_\lambda = (\pm W_\lambda, 0)$, where

$$W(x) = \Big(1 + \frac{|x|^2}{N(N-2)}\Big)^{-(N-2)/2}.$$

The functions W_λ are called *ground states.* They are the only radially symmetric solutions and, up to translation, the only positive solutions of the critical elliptic problem

$$-\Delta u - f(u) = 0. \tag{1.1}$$

Note however that classification of nonradial solutions of (1.1) is an open problem (see for example [73] for details).

Recall that the *Soliton Resolution Conjecture* predicts that a generic bounded (in a suitable sense) solution of a hamiltonian system asymptotically decomposes as a sum of decoupled solitons and a dispersion. This belief is based mainly on the analysis of completely integrable systems, for instance Eckhaus and Schuur [28]. The only complete classification of the dynamical behaviour of a non-integrable hamiltonian system is the result of Duyckaerts, Kenig and Merle [26], which we recall here for the reader's convenience:

Theorem ([26])**.** *Let $N = 3$ and let $\boldsymbol{u}(t) : [t_0, T_+) \to \mathcal{E}$ be a radial solution of* (NLW)*. Then one of the following holds:*

- ***Type I blow-up:*** *$T_+ < \infty$ and*
$$\lim_{t\to T_+} \|\boldsymbol{u}(t)\|_{\mathcal{E}} = +\infty.$$
- ***Type II blow-up:*** *$T_+ < \infty$ and there exist $\boldsymbol{v}_0 \in \mathcal{E}$, an integer $n \in \mathbb{N} \setminus \{0\}$, and for all $j \in \{1, \ldots, n\}$, a sign $\iota_j \in \{\pm 1\}$, and a positive function $\lambda_j(t)$ defined for t close to T_+ such that*
$$\lambda_1(t) \ll \lambda_2(t) \ll \ldots \ll \lambda_n(t) \ll T_+ - t \text{ as } t \to T_+$$
$$\lim_{t\to T_+} \Big\|\boldsymbol{u}(t) - \Big(\boldsymbol{v}_0 + \sum_{j=1}^{n} \iota_j \boldsymbol{W}_{\lambda_j(t)}\Big)\Big\|_{\mathcal{E}} = 0. \tag{1.2}$$
- ***Global solution:*** *$T_+ = +\infty$ and there exist a solution $\boldsymbol{v}_{\text{LIN}}$ of the linear wave equation, an integer $n \in \mathbb{N}$, and for all $j \in \{1, \ldots, n\}$, a sign $\iota_j \in \{\pm 1\}$, and a positive function $\lambda_j(t)$ defined for large t such that*
$$\lambda_1(t) \ll \lambda_2(t) \ll \ldots \ll \lambda_n(t) \ll t \text{ as } t \to +\infty$$
$$\lim_{t\to +\infty} \Big\|\boldsymbol{u}(t) - \Big(\boldsymbol{v}_{\text{LIN}}(t) + \sum_{j=1}^{n} \iota_j \boldsymbol{W}_{\lambda_j(t)}\Big)\Big\|_{\mathcal{E}} = 0. \tag{1.3}$$

□

Of special interest are the solutions which are bounded in $\mathcal{E}$ and which exhibit *no dispersion* (that is, $\boldsymbol{v}_0 = 0$ or $\boldsymbol{v}_{\text{LIN}} = 0$) in one or both time directions. One of the consequences of the *energy channel estimates* in [26] is that in the case $N = 3$ the only solutions without any dispersion in both time directions are the stationary states $\boldsymbol{W}_\lambda$. This is in contrast with the case of completely integrable systems.

In the present paper we are interested in solutions with no dispersion in one time direction, say for positive times. According to Theorem 1.1, for $N = 3$ such a solution has to behave asymptotically as a decoupled superposition of stationary states. Such solutions are called (pure) *multi-bubbles* (or n-bubbles, where n is the number of bubbles). By conservation of energy, if $\boldsymbol{u}(t)$ is an n-bubble, then
$$E(\boldsymbol{u}(t)) = nE(\boldsymbol{W}).$$
The case $n = 1$ in dimension $N \in \{3, 4, 5\}$ was treated by Duyckaerts and Merle [27], who obtained a complete classification of solutions of (NLW) at energy level $E(\boldsymbol{u}(t)) = E(\boldsymbol{W})$. In particular, the only 1-bubbles are $\boldsymbol{W}_\lambda$, $\boldsymbol{W}^-_\lambda$ and $\boldsymbol{W}^+_\lambda$, where $\boldsymbol{W}^-$ and $\boldsymbol{W}^+$ are some special solutions converging exponentially to $\boldsymbol{W}$. The authors solve also the *reconnection problem* by showing that for negative times $\boldsymbol{W}^-$ scatters and $\boldsymbol{W}^+$ blows up in norm $\mathcal{E}$ in finite time.

Solutions of (NLW) satisfying (1.2) or (1.3) with $\boldsymbol{v}_0 \neq 0$ or $\boldsymbol{v}_{\text{LIN}} \neq 0$ can exhibit non-trivial dynamical behaviour, see the results of Krieger, Schlag and Tataru [53], Hillairet and Raphaël [36], Donninger and Krieger [20], Donninger, Huang, Krieger and Schlag [19] and the author [40].

In the present paper we address the case $n = 2$, and more specifically the situation when the two bubbles have opposite signs.

Theorem 1. *Let $N \geq 3$. There exists no radial solutions $\boldsymbol{u} : [t_0, T_+) \to \mathcal{E}$ of* (NLW) *such that*

$$\lim_{t \to T_+} \|\boldsymbol{u}(t) - \boldsymbol{W}_{\lambda_1(t)} + \boldsymbol{W}_{\lambda_2(t)}\|_{\mathcal{E}} = 0 \tag{1.4}$$

and

- *in the case $T_+ < +\infty$, $\lambda_1(t) \ll \lambda_2(t) \ll T_+ - t$ as $t \to T_+$,*
- *in the case $T_+ = +\infty$, $\lambda_1(t) \ll \lambda_2(t) \ll t$ as $t \to +\infty$.*

Remark 1.1. There exist no examples of solutions of (NLW) such that expansion (1.2) or (1.3) holds with $n > 1$ (with or without dispersion). Note however that spatially decoupled non-radial multi-bubbles were recently constructed by Martel and Merle [58] using the Lorentz transform. In their setting, the choice of signs seems to have little importance.

On the other side, Theorem 1 is, to my knowledge, the only result proving non-existence of solutions of type multi-bubble for (NLW) in some specific cases. Existence of pure two-bubbles with the same sign is an open problem.

Remark 1.2. In the case of corotational wave maps existence of pure two-bubbles with the same orientation is easily excluded for variational reasons. Our proof might be seen as an adaptation of this argument to the case where the energy functional is not coercive.

Note that for corotational wave maps existence of pure two-bubbles with opposite orientations is an open problem, related to the *threshold conjecture* for degree 0 equivariant wave maps, see Côte, Kenig, Lawrie and Schlag [16].

Remark 1.3. For the corresponding slightly sub-critical elliptic problem positive multi-bubbles cannot form, whereas multi-bubbles with alternating signs exist, see Li [55], Pistoia and Weth [77].

1.2 Outline of the proof

Step 1. The linearization of (NLW) around $\boldsymbol{W}_\lambda$ has a stable direction $\mathcal{Y}_\lambda^-$. We construct *stable manifolds* $\boldsymbol{U}_\lambda^a$ which are forward invariant sets tangent to $\mathcal{Y}_\lambda^-$ at $\boldsymbol{W}_\lambda$. They have good regularity and decay properties. They allow to define the *refined unstable mode* $\beta_\lambda^a \in \mathcal{E}^*$ with the following crucial property.

Decompose any initial data close to the family of stationary states as $\boldsymbol{u}_0 = \boldsymbol{U}_\lambda^a + \boldsymbol{g}$, with $\boldsymbol{g}$ satisfying natural orthogonality conditions by an appropriate choice of λ and a. We have the alternative:

- (Coercivity) $|\langle \beta_\lambda^a, \boldsymbol{g}\rangle| \lesssim \|\boldsymbol{g}\|_{\mathcal{E}}^2$, which implies $E(\boldsymbol{u}_0) - E(\boldsymbol{W}) \gtrsim \|\boldsymbol{g}\|_{\mathcal{E}}^2$,
- (Destabilization) $|\langle \beta_\lambda^a, \boldsymbol{g}\rangle| \gg \|\boldsymbol{g}\|_{\mathcal{E}}^2$, which implies the exponential growth of $|\langle \beta_\lambda^a, \boldsymbol{g}\rangle|$.

In other words, β_λ^a provides an explicit way of controlling how solutions which violate the coercivity of energy leave a neighbourhood of the stationary states for positive times.

Step 2. Let $\boldsymbol{u}(t) : [t_0; T_+) \to \mathcal{E}$ be a solution of (NLW) which satisfies (1.4). As already mentioned, this implies that

$$E(\boldsymbol{u}) = 2E(\boldsymbol{W}). \tag{1.5}$$

We decompose for any $t \in [t_0, T_+)$:

$$\boldsymbol{u}(t) = \boldsymbol{U}^{a_2(t)}_{\lambda_2(t)} - \boldsymbol{W}_{\lambda_1(t)} + \boldsymbol{g}(t), \qquad \lambda_1(t) \ll \lambda_2(t),$$

with $\boldsymbol{g}(t)$ satisfying natural orthogonality conditions (in fact we use a suitable localization of $\boldsymbol{W}_{\lambda_1(t)}$). From the Taylor formula we obtain

$$E(\boldsymbol{u}) = E(\boldsymbol{U}^{a_2}_{\lambda_2} - \boldsymbol{W}_{\lambda_1}) + \langle \mathrm{D}E(\boldsymbol{U}^{a_2}_{\lambda_2} - \boldsymbol{W}_{\lambda_1}, \boldsymbol{g}\rangle + \frac{1}{2}\langle \mathrm{D}^2 E(\boldsymbol{U}^{a_2}_{\lambda_2} - \boldsymbol{W}_{\lambda_1})\boldsymbol{g}, \boldsymbol{g}\rangle + o(\|\boldsymbol{g}\|_{\mathcal{E}}^2). \tag{1.6}$$

An explicit key computation shows that

$$E(\boldsymbol{U}^{a_2}_{\lambda_2} - \boldsymbol{W}_{\lambda_1}) - 2E(\boldsymbol{W}) \gtrsim (\lambda_1/\lambda_2)^{\frac{N-2}{2}}.$$

It is at this point that the sign condition is decisive.

Step 3. We prove that the assumption that $\boldsymbol{u}(t)$ stays close to a 2-bubble implies that $|\langle \beta^{a_2}_{\lambda_2}, \boldsymbol{g}\rangle| \lesssim \|\boldsymbol{g}\|_{\mathcal{E}}^2 + (\lambda_1/\lambda_2)^{\frac{N-2}{2}}$. This allows to show that the second term in the expansion (1.6) is $\ll \|\boldsymbol{g}\|_{\mathcal{E}}^2 + (\lambda_1/\lambda_2)^{\frac{N-2}{2}}$.

Finally, by an elementary analysis of the linear stable and unstable modes we can prove that, at least along a sequence of times, the third term of the expansion (1.6) is coercive, that is $\gtrsim \|\boldsymbol{g}\|_{\mathcal{E}}^2$. Inserted in (1.6), this leads to $E(\boldsymbol{u}) > 2E(\boldsymbol{W})$, contradicting (1.5).

1.3 Notation

We introduce the infinitesimal generators of the scale change

$$\Lambda_s := \Big(\frac{N}{2} - s\Big) + x \cdot \nabla.$$

For $s = 1$ we omit the subscript and write $\Lambda = \Lambda_1$. We denote $\Lambda_{\mathcal{E}}$, $\Lambda_{\mathcal{F}}$ and $\Lambda_{\mathcal{E}^*}$ the infinitesimal generators of the scaling which is critical for a given norm, that is

$$\Lambda_{\mathcal{E}} = (\Lambda, \Lambda_0), \quad \Lambda_{\mathcal{F}} = (\Lambda_0, \Lambda_{-1}), \quad \Lambda_{\mathcal{E}^*} = (\Lambda_{-1}, \Lambda_0).$$

We use the subscript $\cdot_\lambda$ to denote rescaling with characteristic length λ, critical for a norm which will be known from the context.

We introduce the following notation for some frequently used function spaces: $X^s := \dot H^{s+1}_{\mathrm{rad}} \cap \dot H^1_{\mathrm{rad}}$ for $s \geq 0$, $Y^k := H^k(1 + |x|^k)$ for $k \in \mathbb{N}$, $\mathcal{E} := \dot H^1_{\mathrm{rad}} \times L^2_{\mathrm{rad}}$, $\mathcal{F} := L^2_{\mathrm{rad}} \times \dot H^{-1}_{\mathrm{rad}}$. Notice that $\mathcal{E}^* \simeq \dot H^{-1}_{\mathrm{rad}} \times L^2_{\mathrm{rad}}$ through the natural isomorphism given by the distributional pairing. In the sequel we will omit the subscript and write $\dot H^1$ for $\dot H^1_{\mathrm{rad}}$ etc. We denote $J := \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0\end{pmatrix}$; note that $J\mathcal{E}^* = \mathcal{F}$.

For a function space $\mathcal{A}$, $O_{\mathcal{A}}(m)$ denotes any $a \in \mathcal{A}$ such that $\|a\|_{\mathcal{A}} \leq Cm$ for some constant $C > 0$. We denote $B_{\mathcal{A}}(x_0, \eta)$ an open ball of center x_0 and radius η in the space $\mathcal{A}$. If $\mathcal{A}$ is not specified, it means that $\mathcal{A} = \mathbb{R}$.

For a radial function $g : \mathbb{R}^N \to \mathbb{R}$ and $r \geq 0$ we denote $g(r)$ the value of $g(x)$ for $|x| = r$.

2 Sharp coercivity properties near $\boldsymbol{W}_\lambda$

2.1 Properties of the linearized operator

Linearizing (NLW) around $\boldsymbol{W}$, $\boldsymbol{u} = \boldsymbol{W} + \boldsymbol{g}$, one obtains

$$\partial_t \boldsymbol{g} = J \circ \mathrm{D}^2 E(\boldsymbol{W})\boldsymbol{g} = \begin{pmatrix} 0 & \mathrm{Id} \\ -L & 0 \end{pmatrix} \boldsymbol{g},$$

where L is the Schrödinger operator

$$Lg := (-\Delta - f'(W))g.$$

Notice that $L(\Lambda W) = \frac{\mathrm{d}}{\mathrm{d}\lambda}\big|_{\lambda=1}\big(- \Delta W_\lambda - f(W_\lambda)\big) = 0$. It is known that L has exactly one strictly negative simple eigenvalue which we denote $-\nu^2$ (we take $\nu > 0$). We denote the corresponding positive eigenfunction $\mathcal{Y}$, normalized so that $\|\mathcal{Y}\|_{L^2} = 1$. By elliptic regularity $\mathcal{Y}$ is smooth and by Agmon estimates it decays exponentially. Self-adjointness of L implies that

$$\langle \mathcal{Y}, \Lambda W\rangle = 0.$$

Fix $\mathcal{Z} \in C_0^\infty$ such that

$$\langle \mathcal{Z}, \Lambda W\rangle > 0, \qquad \langle \mathcal{Z}, \mathcal{Y}\rangle = 0.$$

We have the following linear (localized) coercivity result, similar to [58, Lemma 2.1].

Lemma 2.1. *There exist constants $c, C > 0$ such that*

- *for all $g \in \dot H^1$ radially symmetric there holds*

$$\langle g, Lg\rangle = \int_{\mathbb{R}^N} |\nabla g|^2\,\mathrm{d}x - \int_{\mathbb{R}^N} f'(W)|g|^2\,\mathrm{d}x \geq c\int_{\mathbb{R}^N}|\nabla g|^2\,\mathrm{d}x - C\big(\langle \mathcal{Z}, g\rangle^2 + \langle \mathcal{Y}, g\rangle^2\big), \tag{2.1}$$

- *if $r_1 > 0$ is large enough, then for all $g \in \dot H^1_{\mathrm{rad}}$ there holds*

$$(1-2c)\int_{|x|\leq r_1}|\nabla g|^2\,\mathrm{d}x + c\int_{|x|\geq r_1}|\nabla g|^2\,\mathrm{d}x - \int_{\mathbb{R}^N} f'(W)|g|^2\,\mathrm{d}x \geq -C\big(\langle \mathcal{Z}, g\rangle^2 + \langle \mathcal{Y}, g\rangle^2\big), \tag{2.2}$$

- *if $r_2 > 0$ is small enough, then for all $g \in \dot H^1_{\mathrm{rad}}$ there holds*

$$(1-2c)\int_{|x|\geq r_2}|\nabla g|^2\,\mathrm{d}x + c\int_{|x|\leq r_2}|\nabla g|^2\,\mathrm{d}x - \int_{\mathbb{R}^N} f'(W)|g|^2\,\mathrm{d}x \geq -C\big(\langle \mathcal{Z}, g\rangle^2 + \langle \mathcal{Y}, g\rangle^2\big). \tag{2.3}$$

Proof. We will prove (2.2) and (2.3). For a proof of (2.1) we refer to [40, Lemma 6.1], see also [27, Proposition 5.5] for a different formulation.

We define the projections $\Pi_r, \Psi_r : \dot H^1 \to \dot H^1$:

$$(\Pi_r g)(x) := \begin{cases} g(r) & \text{if } |x| \leq r, \\ g(x) & \text{if } |x| \geq r, \end{cases} \qquad (\Psi_r g)(x) := \begin{cases} g(x) - g(r) & \text{if } |x| \leq r, \\ 0 & \text{if } |x| \geq r \end{cases}$$

(thus $\Pi_r + \Psi_r = \mathrm{Id}$).

Applying (2.1) to $\Psi_{r_1} g$ with c replaced by $3c$ and C replaced by $\frac{C}{2}$ we get

$$(1-2c)\int_{|x|\leq r_1} |\nabla g|^2 \,\mathrm{d}x = (1-2c)\int_{\mathbb{R}^N} |\nabla(\Psi_{r_1} g)|^2 \,\mathrm{d}x \\ \geq (1+c)\int_{\mathbb{R}^N} f'(W)|\Psi_{r_1} g|^2 \,\mathrm{d}x - \frac{C}{2}\big(\langle \mathcal{Z}, \Psi_{r_1} g\rangle^2 + \langle \mathcal{Y}, \Psi_{r_1} g\rangle^2\big). \tag{2.4}$$

By Sobolev and Hölder inequalities we have

$$\int_{|x|\geq r_1} f'(W)|g|^2 \,\mathrm{d}x = \int_{|x|\geq r_1} f'(W)|\Pi_{r_1} g|^2 \,\mathrm{d}x \\ \lesssim \|f'(W)\|_{L^{\frac{N}{2}}(|x|\geq r_1)} \cdot \|\Pi_{r_1} g\|^2_{\dot{H}^1} \leq \frac{c}{4}\int_{|x|\geq r_1} |\nabla g|^2 \,\mathrm{d}x \tag{2.5}$$

if r_1 is large enough.

In the region $|x| \leq r_1$ we apply the pointwise inequality

$$|g(x)|^2 \leq (1+c)|(\Psi_{r_1} g)(x)|^2 + (1+c^{-1})|g(r_1)|^2, \qquad |x| \leq r_1. \tag{2.6}$$

Recall that by the Strauss Lemma [87], for a radial function g there holds

$$|g(r_1)| \lesssim \|\Pi_{r_1} g\|_{\dot{H}^1} \cdot r_1^{-\frac{N-2}{2}}.$$

Since $f'(W(r)) \sim r^{-4}$ as $r \to +\infty$, we have

$$\int_{|x|\leq r_1} f'(W) \,\mathrm{d}x \ll r_1^{N-2}, \qquad \text{as } r_1 \to +\infty,$$

hence

$$\int_{|x|\leq r_1} f'(W)\cdot(1+c^{-1})|g(r_1)|^2 \,\mathrm{d}x \leq \frac{c}{4}\int_{|x|\geq r_1} |\nabla g|^2 \,\mathrm{d}x \tag{2.7}$$

if r_1 is large enough.

Estimates (2.5), (2.6) and (2.7) yield

$$\int_{\mathbb{R}^N} f'(W)|g|^2 \,\mathrm{d}x \leq (1+c)\int_{\mathbb{R}^N} f'(W)|(\Psi_{r_1} g)(x)|^2 \,\mathrm{d}x + \frac{c}{2}\int_{|x|\geq r_1} |\nabla g|^2 \,\mathrm{d}x. \tag{2.8}$$

Using the fact that $\mathcal{Y} \in L^1 \cap L^{\frac{2N}{N+2}}$ we obtain

$$|\langle \mathcal{Y}, \Pi_{r_1} g\rangle| \lesssim \|\Pi_{r_1} g\|_{\dot{H}^1} \cdot r_1^{-\frac{N-2}{2}} + \int_{|x|\geq r_1} \mathcal{Y}|g| \,\mathrm{d}x \lesssim (r_1^{-\frac{N-2}{2}} + \|\mathcal{Y}\|_{L^{\frac{2N}{N+2}}(|x|\geq r_1)})\|\Pi_{r_1} g\|_{\dot{H}^1},$$

hence

$$\frac{C}{2}\langle \mathcal{Y}, \Psi_{r_1} g\rangle^2 \leq C\langle \mathcal{Y}, g\rangle^2 + C\langle \mathcal{Y}, \Pi_{r_1} g\rangle^2 \leq C\langle \mathcal{Y}, g\rangle^2 + \frac{c}{4}\int_{|x|\geq r_1} |\nabla g|^2 \,\mathrm{d}x, \tag{2.9}$$

provided that r_1 is chosen large enough. Similarly,

$$\frac{C}{2}\langle \mathcal{Z}, \Psi_{r_1} g\rangle^2 \leq C\langle \mathcal{Z}, g\rangle^2 + C\langle \mathcal{Z}, \Pi_{r_1} g\rangle^2 \leq C\langle \mathcal{Z}, g\rangle^2 + \frac{c}{4}\int_{|x|\geq r_1} |\nabla g|^2 \,\mathrm{d}x. \tag{2.10}$$

Estimate (2.2) follows from (2.4), (2.8), (2.9) and (2.10).

We turn to the proof of (2.3). Applying (2.1) to $\Pi_{r_2}g$ with c replaced by $3c$ and C replaced by $\frac{C}{2}$ we get

$$\begin{aligned}(1-3c)\int_{|x|\geq r_2}|\nabla g|^2\,\mathrm{d}x &= (1-3c)\int_{\mathbb{R}^N}|\nabla(\Pi_{r_2}g)|^2\,\mathrm{d}x \\ &\geq \int_{\mathbb{R}^N} f'(W)|\Pi_{r_2}g|^2\,\mathrm{d}x - \frac{C}{2}\big(\langle\mathcal{Z},\Pi_{r_2}g\rangle^2+\langle\mathcal{Y},\Pi_{r_2}g\rangle^2\big).\end{aligned}\tag{2.11}$$

By Sobolev and Hölder inequalities we have for r_2 small enough

$$\int_{|x|\leq r_2} f'(W)|g|^2\,\mathrm{d}x \leq \frac{c}{2}\int_{\mathbb{R}^N}|\nabla g|^2\,\mathrm{d}x. \tag{2.12}$$

By definition of Π_r there holds

$$\int_{|x|\geq r_2} f'(W)|g|^2\,\mathrm{d}x \leq \int_{\mathbb{R}^N} f'(W)|\Pi_{r_2}g|^2\,\mathrm{d}x,$$

hence (2.11) and (2.12) imply

$$(1-2c)\int_{|x|\geq r_2}|\nabla g|^2\,\mathrm{d}x+\frac{c}{2}\int_{|x|\leq r_2}|\nabla g|^2\,\mathrm{d}x \geq \int_{\mathbb{R}^N} f'(W)|g|^2\,\mathrm{d}x-\frac{C}{2}\big(\langle\mathcal{Z},\Pi_{r_2}g\rangle^2+\langle\mathcal{Y},\Pi_{r_2}g\rangle^2\big). \tag{2.13}$$

Using the fact that $\mathcal{Y}\in L^{\frac{2N}{N+2}}$ we obtain

$$|\langle\mathcal{Y},\Psi_{r_2}g\rangle| \lesssim \int_{|x|\leq r_2}\mathcal{Y}|g|\,\mathrm{d}x \lesssim \|\mathcal{Y}\|_{L^{\frac{2N}{N+2}}(|x|\leq r_2)}\|\Psi_{r_2}g\|_{\dot H^1},$$

hence

$$\frac{C}{2}\langle\mathcal{Y},\Pi_{r_2}g\rangle^2 \leq C\langle\mathcal{Y},g\rangle^2+C\langle\mathcal{Y},\Psi_{r_2}g\rangle^2 \leq C\langle\mathcal{Y},g\rangle^2+\frac{c}{4}\int_{|x|\leq r_2}|\nabla g|^2\,\mathrm{d}x, \tag{2.14}$$

provided that r_2 is chosen small enough. Similarly,

$$\frac{C}{2}\langle\mathcal{Z},\Pi_{r_2}g\rangle^2 \leq C\langle\mathcal{Z},g\rangle^2+C\langle\mathcal{Z},\Psi_{r_2}g\rangle^2 \leq C\langle\mathcal{Z},g\rangle^2+\frac{c}{4}\int_{|x|\leq r_2}|\nabla g|^2\,\mathrm{d}x. \tag{2.15}$$

Estimate (2.3) follows from (2.13), (2.14) and (2.15).

□

We define

$$\mathcal{Y}^- := \big(\frac{1}{\nu}\mathcal{Y},-\mathcal{Y}\big),\qquad \mathcal{Y}^+ := \big(\frac{1}{\nu}\mathcal{Y},\mathcal{Y}\big),$$
$$\alpha^- := \frac{\nu}{2}J\mathcal{Y}^+ = \frac{1}{2}(\nu\mathcal{Y},-\mathcal{Y}),\qquad \alpha^+ := -\frac{\nu}{2}J\mathcal{Y}^- = \frac{1}{2}(\nu\mathcal{Y},\mathcal{Y}).$$

We have $J\circ\mathrm{D}^2E(\boldsymbol{W}) = \begin{pmatrix}0 & \mathrm{Id}\\ -L & 0\end{pmatrix}$. A short computation shows that

$$J\circ\mathrm{D}^2E(\boldsymbol{W})\mathcal{Y}^- = -\nu\mathcal{Y}^-,\qquad J\circ\mathrm{D}^2E(\boldsymbol{W})\mathcal{Y}^+ = \nu\mathcal{Y}^+$$

and

$$\langle\alpha^-,J\circ\mathrm{D}^2E(\boldsymbol{W})\boldsymbol{g}\rangle = -\nu\langle\alpha^-,\boldsymbol{g}\rangle,\qquad \langle\alpha^+,J\circ\mathrm{D}^2E(\boldsymbol{W})\boldsymbol{g}\rangle = \nu\langle\alpha^+,\boldsymbol{g}\rangle,\qquad \forall\boldsymbol{g}\in\mathcal{E}. \tag{2.16}$$

We will think of α^- and α^+ as linear forms on $\mathcal{E}$. Notice that $\langle\alpha^-,\mathcal{Y}^-\rangle = \langle\alpha^+,\mathcal{Y}^+\rangle = 1$ and $\langle\alpha^-,\mathcal{Y}^+\rangle = \langle\alpha^+,\mathcal{Y}^-\rangle = 0$.

The rescaled versions of these objects are

$$\mathcal{Y}_\lambda^- := \big(\frac{1}{\nu}\mathcal{Y}_\lambda, -\mathcal{Y}_{\underline{\lambda}}\big), \qquad \mathcal{Y}_\lambda^+ := \big(\frac{1}{\nu}\mathcal{Y}_\lambda, \mathcal{Y}_{\underline{\lambda}}\big),$$
$$\alpha_\lambda^- := \frac{\nu}{2\lambda}J\mathcal{Y}_\lambda^+ = \frac{1}{2}\big(\frac{\nu}{\lambda}\mathcal{Y}_{\underline{\lambda}}, -\mathcal{Y}_{\underline{\lambda}}\big), \qquad \alpha_\lambda^+ := -\frac{\nu}{2\lambda}J\mathcal{Y}_\lambda^- = \frac{1}{2}\big(\frac{\nu}{\lambda}\mathcal{Y}_{\underline{\lambda}}, \mathcal{Y}_{\underline{\lambda}}\big). \tag{2.17}$$

The scaling is chosen so that $\langle\alpha_\lambda^-,\mathcal{Y}_\lambda^-\rangle = \langle\alpha_\lambda^+,\mathcal{Y}_\lambda^+\rangle = 1$. We have

$$J\circ \mathrm{D}^2E(\boldsymbol{W}_\lambda)\mathcal{Y}_\lambda^- = -\frac{\nu}{\lambda}\mathcal{Y}_\lambda^-, \qquad J\circ \mathrm{D}^2E(\boldsymbol{W}_\lambda)\mathcal{Y}_\lambda^+ = \frac{\nu}{\lambda}\mathcal{Y}_\lambda^+ \tag{2.18}$$

and

$$\langle\alpha_\lambda^-, J\circ \mathrm{D}^2E(\boldsymbol{W}_\lambda)\boldsymbol{g}\rangle = -\frac{\nu}{\lambda}\langle\alpha_\lambda^-,\boldsymbol{g}\rangle, \qquad \langle\alpha_\lambda^+, J\circ \mathrm{D}^2E(\boldsymbol{W}_\lambda)\boldsymbol{g}\rangle = \frac{\nu}{\lambda}\langle\alpha_\lambda^+,\boldsymbol{g}\rangle, \qquad \forall \boldsymbol{g}\in\mathcal{E}. \tag{2.19}$$

As a standard consequence of (2.1), we obtain the following:

Lemma 2.2. *There exists a constant $\eta > 0$ such that if $\|\boldsymbol{V}-\boldsymbol{W}_\lambda\|_\mathcal{E} < \eta$, then for all $\boldsymbol{g}\in\mathcal{E}$ such that $\langle\mathcal{Z}_{\underline{\lambda}}, g\rangle = 0$ there holds*

$$\frac{1}{2}\langle \mathrm{D}^2E(\boldsymbol{V})\boldsymbol{g},\boldsymbol{g}\rangle + 2\big(\langle\alpha_\lambda^-,\boldsymbol{g}\rangle^2 + \langle\alpha_\lambda^+,\boldsymbol{g}\rangle^2\big) \gtrsim \|\boldsymbol{g}\|_\mathcal{E}^2.$$

Proof. For $N\in\{3,4,5\}$ see [39, Lemma 2.2]. For $N\geq 6$ the same proof is valid, once we notice that $\|f'(V)-f'(W_\lambda)\|_{L^{\frac{N}{2}}} \leq f'(\|V-W_\lambda\|_{\dot H^1})$. □

We now turn to the proofs of various energy estimates for the linear group generated by

$$A := J\circ \mathrm{D}^2E(\boldsymbol{W}) = \begin{pmatrix} 0 & \mathrm{Id} \\ -L & 0\end{pmatrix}.$$

on its invariant subspaces, which will be needed in Subsection 2.2. This is much in the spirit of [5, Section 2].

It follows from (2.16) that the *centre-stable subspace* $\mathcal{X}_{\mathrm{cs}} := \ker\alpha^+$, the *centre-unstable subspace* $\mathcal{X}_{\mathrm{cu}} := \ker\alpha^-$ and the *centre subspace* $\mathcal{X}_{\mathrm{c}} := \mathcal{X}_{\mathrm{cs}}\cap\mathcal{X}_{\mathrm{cu}}$ are invariant subspaces of the operator A. Notice that $\langle\alpha^-,\mathcal{Y}^-\rangle = \langle\alpha^+,\mathcal{Y}^+\rangle = 1$, $\mathcal{E} = \mathcal{X}_{\mathrm{cs}}\oplus\{a\mathcal{Y}^+\} = \mathcal{X}_{\mathrm{cu}}\oplus\{a\mathcal{Y}^-\}$, $\mathcal{X}_{\mathrm{cs}} = \mathcal{X}_{\mathrm{c}}\oplus\{a\mathcal{Y}^-\}$, $\mathcal{X}_{\mathrm{cu}} = \mathcal{X}_{\mathrm{c}}\oplus\{a\mathcal{Y}^+\}$.

We define $\mathcal{X}_{\mathrm{cc}} := \{\boldsymbol{v} = (v,\dot v)\in\mathcal{X}_{\mathrm{c}} \mid \langle\mathcal{Z}, v\rangle = 0\}$.

Lemma 2.3. *Let $k\in\mathbb{N}$. There exist constants $1 = a_0 > a_1 > \ldots > a_k > 0$ such that the norm $\|\cdot\|_{A,k}$ defined by the following formula:*

$$\|\boldsymbol{v}\|_{A,k}^2 := \sum_{j=0}^k a_j\big(\langle v, L^{j+1}v\rangle + \langle\dot v, L^j\dot v\rangle\big)$$

satisfies $\|\boldsymbol{v}\|_{X^k\times H^k} \sim \|\boldsymbol{v}\|_{A,k}$ for all $\boldsymbol{v} = (v,\dot v)\in(X^k\times H^k)\cap\mathcal{X}_{\mathrm{cc}}$.

Proof. We proceed by induction. For $k=0$ we have

$$\|\boldsymbol{v}\|_{A,0} = \sqrt{\langle v, Lv\rangle + \langle\dot v,\dot v\rangle} = \sqrt{\langle \mathrm{D}^2E(\boldsymbol{W})\boldsymbol{v},\boldsymbol{v}\rangle}.$$

By Lemma 2.2, this norm is equivalent to $\|\cdot\|_\mathcal{E}$ on $\mathcal{E}\cap\mathcal{X}_{\mathrm{cc}}$.

To check the induction step, one should show that for any $k > 0$ there exist $a_1, a_2 > 0$ such that

$$\|v\|_1^2 := \|v\|_{X^{k-1}}^2 + a_1\langle v, L^{k+1}v\rangle \gtrsim \|v\|_{X^k}^2 \tag{2.20}$$

and

$$\|\dot v\|_2^2 := \|\dot v\|_{H^{k-1}}^2 + a_2\langle \dot v, L^k \dot v\rangle \gtrsim \|\dot v\|_{H^k}^2. \tag{2.21}$$

To prove (2.21) notice that

$$L^k = (-\Delta)^k + (\text{terms with at most } 2k-2 \text{ derivatives}).$$

Integrating by parts all the terms except for the first one we arrive at expressions of the form $\int V \cdot \partial^i v \cdot \partial^j v \,\mathrm{d}x$ where V is bounded and $i, j \leq k-1$. All these expressions are controlled by $\|\cdot\|_{H^{k-1}}^2$.

The proof of (2.20) is almost the same. The only problem are the terms of the form $\int V \cdot \nabla v \cdot v \,\mathrm{d}x$ and $\int V \cdot |v|^2 \,\mathrm{d}x$. As the potential decreases at least as $f'(W)$, by Hardy inequality these terms are controlled by $\|v\|_{\dot H^1}$. □

We will denote $\langle\cdot,\cdot\rangle_{A,k}$ the scalar product associated with the norm $\|\cdot\|_{A,k}$.

We define the projections:

$$\pi_s \boldsymbol{v} := \langle \alpha^-, \boldsymbol{v}\rangle \mathcal{Y}^-, \qquad \pi_{cu} := \mathrm{Id} - \pi_s.$$

We denote π_{cc} the projection of $\mathcal{X}_\mathrm{c}$ on $\mathcal{X}_\mathrm{cc}$ in the direction $\Lambda_\mathcal{E}\boldsymbol{W}$. These projections are continuous linear operators on $\mathcal{E}$ as well as on $X^k \times H^k$ for $k > 0$.

Proposition 2.4. *The operator A generates a strongly continuous group on $X^k \times H^k$ denoted e^{tA}. Moreover, the following bounds are true for $t \geq 0$:*

$$\boldsymbol{v}_0 \in (X^k \times H^k) \cap \mathcal{X}_\mathrm{s} \quad \Rightarrow \quad \|\mathrm{e}^{tA}\boldsymbol{v}_0\|_{X^k\times H^k} \lesssim \mathrm{e}^{-\nu t}\|\boldsymbol{v}_0\|_{X^k\times H^k}, \tag{2.22}$$

$$\boldsymbol{v}_0 \in (X^k \times H^k) \cap \mathcal{X}_\mathrm{cu} \quad \Rightarrow \quad \|\mathrm{e}^{-tA}\boldsymbol{v}_0\|_{X^k\times H^k} \lesssim (1+t)\|\boldsymbol{v}_0\|_{X^k\times H^k}, \tag{2.23}$$

$$\boldsymbol{v}_0 \in X^k \times H^k \quad \Rightarrow \quad \|\mathrm{e}^{-tA}\boldsymbol{v}_0\|_{X^k\times H^k} \lesssim \mathrm{e}^{\nu t}\|\boldsymbol{v}_0\|_{X^k\times H^k}. \tag{2.24}$$

Proof. It suffices to analyse the restriction to the invariant subspace $\mathcal{X}_\mathrm{c}$. Take $\boldsymbol{v}_0 \in \mathcal{X}_\mathrm{c}$ and decompose $\boldsymbol{v}_0 = l_0\Lambda_\mathcal{E}\boldsymbol{W} + \boldsymbol{w}_0$, $\boldsymbol{w}_0 \in \mathcal{X}_\mathrm{cc}$ (notice that $\Lambda_\mathcal{E}\boldsymbol{W} \in X^k \times H^k$). It can be checked that the operator $B := \pi_{cc}\circ A$ is skew-adjoint for the scalar product $\langle\cdot,\cdot\rangle_{A,k}$, hence it generates a unitary group $\boldsymbol{w}(t) = \mathrm{e}^{tB}\boldsymbol{w}_0$ by the Stone theorem. Let $l(t)$ be defined by the formula

$$l(t) = l_0 + \int_0^t \frac{\langle \mathcal{Z}, \dot w(t)\rangle}{\langle \mathcal{Z}, \Lambda W\rangle}\,\mathrm{d}t. \tag{2.25}$$

Set $\boldsymbol{v}(t) = \boldsymbol{w}(t) + l(t)\Lambda_\mathcal{E}\boldsymbol{W}$. This defines a linear group and

$$\lim_{t\to 0}\frac{1}{t}(\boldsymbol{v}(t) - \boldsymbol{v}_0) = B\boldsymbol{w}_0 + l'(0)\Lambda_\mathcal{E}\boldsymbol{W} = B\boldsymbol{v}_0 + \frac{\langle \mathcal{Z}, \dot w_0\rangle}{\langle \mathcal{Z}, \Lambda W\rangle}\Lambda_\mathcal{E}\boldsymbol{W} = A\boldsymbol{v}_0,$$

hence $\boldsymbol{v}(t) = \mathrm{e}^{tA}\boldsymbol{v}_0$.

Estimate (2.22) follows immediately from the fact that $\mathcal{Y}^-$ is an eigenfunction of A with eigenvalue $-\nu$. Analogously, in (2.23) we can assume that $\boldsymbol{v}_0 \in \mathcal{X}_\mathrm{c}$ (the unstable mode decreases exponentially for negative times). By the equivalence of norms and the fact that the group generated by B is unitary for the norm $\|\cdot\|_{A,k}$,

$$\|\boldsymbol{w}(t)\|_{X^k\times H^k} \lesssim \|\boldsymbol{v}_0\|_{X^k\times H^k} \quad \text{for all } t, \tag{2.26}$$

hence it suffices to bound $l(t)$. Using (2.26) and the fact that $|l_0| \lesssim \|\boldsymbol{v}_0\|_{X^k\times H^k}$ we get from (2.25) that $|l(t)| \lesssim (1+|t|)\|\boldsymbol{v}_0\|_{X^k\times H^k}$.

Estimate (2.24) follows easily from (2.23). □

Remark 2.5. The factor $|t|$ in (2.23) is necessary, for example in dimension $N = 5$ we have a solution $\boldsymbol{v}(t) = (t\Lambda W, \Lambda W)$.

It is possible to finish the construction for example in the space $X^1 \times H^1$. However, later we will need some information on the spatial decay of the constructed functions, which forces us to use weighted spaces. We define

$$\|v\|_{Y^k} := \|(1 + |x|^k)v\|_{H^k}.$$

One may check by induction on $j = 0, 1, \ldots, k$ that

$$\|(1 + |x|^k)v\|_{H^j}^2 \sim \int (1 + |x|^{2k})(|v|^2 + |\nabla^j v|^2)\,\mathrm{d}x,$$

in particular

$$\|v\|_{Y^k}^2 \sim \int (1 + |x|^{2k})(|v|^2 + |\nabla^k v|^2)\,\mathrm{d}x.$$

Lemma 2.6. *Let $k \in \mathbb{Z}, k \geq 0$. The following bounds are true for $t \geq 0$:*

$$\boldsymbol{v}_0 \in (Y^{k+1} \times Y^k) \cap \mathcal{X}_{\mathrm{s}} \quad \Rightarrow \quad \|\mathrm{e}^{tA}\boldsymbol{v}_0\|_{Y^{k+1}\times Y^k} \lesssim \mathrm{e}^{-\nu t}\|\boldsymbol{v}_0\|_{Y^{k+1}\times Y^k}, \tag{2.27}$$

$$\boldsymbol{v}_0 \in (Y^{k+1} \times Y^k) \cap \mathcal{X}_{\mathrm{cu}} \quad \Rightarrow \quad \|\mathrm{e}^{-tA}\boldsymbol{v}_0\|_{Y^{k+1}\times Y^k} \lesssim (1 + t^{\frac{k(k+1)}{2}+1})\|\boldsymbol{v}_0\|_{Y^{k+1}\times Y^k}, \tag{2.28}$$

$$\boldsymbol{v}_0 \in Y^{k+1} \times Y^k \quad \Rightarrow \quad \|\mathrm{e}^{-tA}\boldsymbol{v}_0\|_{Y^{k+1}\times Y^k} \lesssim \mathrm{e}^{\nu t}\|\boldsymbol{v}_0\|_{Y^{k+1}\times Y^k}. \tag{2.29}$$

Proof. The proof of (2.27) and (2.29) is the same as in Proposition 2.4, once we recall that $\mathcal{Y}^- \in Y^{k+1} \times Y^k$. In order to prove (2.28), write $\boldsymbol{v}(t) = \mathrm{e}^{-tA}\boldsymbol{v}_0$, so that $\partial_t \boldsymbol{v} = -A\boldsymbol{v} = (-\dot{v}, -\Delta v - f'(W)v)$, hence

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int (1 + |x|^{2k})(|\dot{v}|^2 + |\nabla v|^2)\,\mathrm{d}x &= -\int (1 + |x|^{2k})\big((\Delta v + f'(W)v)\cdot \dot{v} + \nabla \dot{v}\cdot \nabla v\big)\,\mathrm{d}x \\ &= \int \nabla(|x|^{2k})\cdot \nabla v \cdot \dot{v} + (1 + |x|^{2k})(f'(W)v)\cdot \dot{v}\,\mathrm{d}x\end{aligned}$$

(we have integrated by parts between the first and the second line). Notice that $xf'(W) \in L^N$, hence by Hölder and Sobolev $\|xf'(W)v\|_{L^2} \lesssim \|v\|_{L^{\frac{2N}{N-2}}} \lesssim \|\nabla v\|_{L^2}$, thus

$$\Big|\frac{\mathrm{d}}{\mathrm{d}t}\int (1 + |x|^{2k})(|\dot{v}|^2 + |\nabla v|^2)\,\mathrm{d}x\Big| \lesssim \int (1 + |x|^{2k-1})(|\dot{v}|^2 + |\nabla v|^2)\,\mathrm{d}x. \tag{2.30}$$

Analogously, from

$$\begin{aligned}&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int (1 + |x|^{2k})(|\nabla^k \dot{v}|^2 + |\nabla^{k+1}v|^2)\,\mathrm{d}x \\ &= -\int (1 + |x|^{2k})\big(\nabla^k(\Delta v + f'(W)v)\cdot \nabla^k \dot{v} + \nabla^{k+1}\dot{v}\cdot \nabla^{k+1}v\big)\,\mathrm{d}x \\ &= \int \nabla(|x|^{2k})\cdot \nabla^{k+1}v\cdot \nabla^k \dot{v} + (1 + |x|^{2k})\nabla^k(f'(W)v)\cdot \nabla^k \dot{v}\,\mathrm{d}x\end{aligned}$$

we deduce

$$\Big|\frac{\mathrm{d}}{\mathrm{d}t}\int (1+|x|^{2k})(|\nabla^k \dot{v}|^2+|\nabla^{k+1}v|^2)\,\mathrm{d}x\Big| \lesssim \int (1+|x|^{2k-1})(|\dot{v}|^2+|\nabla^k \dot{v}|^2+|\nabla v|^2+|\nabla^{k+1}v|^2)\,\mathrm{d}x. \tag{2.31}$$

Using (2.30), (2.31) and Hölder we obtain

$$\Big|\frac{\mathrm{d}}{\mathrm{d}t}\int(1+|x|^{2k})(|\dot v|^2+|\nabla^k\dot v|^2+|\nabla v|^2+|\nabla^{k+1}v|^2)\,\mathrm{d}x\Big|$$
$$\lesssim\Big(\int(1+|x|^{2k})(|\dot v|^2+|\nabla^k\dot v|^2+|\nabla v|^2+|\nabla^{k+1}v|^2)\,\mathrm{d}x\Big)^{\frac{2k-1}{2k}}\cdot\|\boldsymbol v\|_{X^k\times H^k}^{\frac1k},$$

which gives, using (2.23) and integrating,

$$\int(1+|x|^{2k})(|\dot v|^2+|\nabla^k\dot v|^2+|\nabla v|^2+|\nabla^{k+1}v|^2)\,\mathrm{d}x\lesssim(1+t^{k(k+1)})\|\boldsymbol v_0\|^2_{Y^{k+1}\times Y^k}.$$

Now we can easily bound the L^2 term by Schwarz inequality:

$$\Big|\frac{\mathrm{d}}{\mathrm{d}t}\int(1+|x|^{2k})|v|^2\,\mathrm{d}x\Big|\lesssim\Big(\int(1+|x|^{2k})|v|^2\,\mathrm{d}x\Big)^{\frac12}\cdot\Big(\int(1+|x|^{2k})|\dot v|^2\,\mathrm{d}x\Big)^{\frac12},$$

which leads to

$$\int(1+|x|^{2k})|v|^2\,\mathrm{d}x\lesssim(1+t^{k(k+1)+2})\|\boldsymbol v_0\|^2_{Y^{k+1}\times Y^k}.$$

□

We fix $k\in\mathbb N$ large enough. For $\widetilde\nu>0$ the space $BC_{\widetilde\nu}$ is defined as the space of continuous functions $\boldsymbol v:[0,+\infty)\to Y^{k+1}\times Y^k$ with the norm

$$\|\boldsymbol v\|_{BC_{\widetilde\nu}}:=\sup_{t\in[0,+\infty)}\mathrm{e}^{\widetilde\nu t}\|\boldsymbol v(t)\|_{Y^{k+1}\times Y^k}.$$

Lemma 2.7. *If $\widetilde\nu\in(0,\nu)$, then for any $\boldsymbol w\in BC_{\widetilde\nu}$ the equation*

$$\partial_t\boldsymbol v(t)=A\boldsymbol v(t)+\boldsymbol w(t)\tag{2.32}$$

has a unique solution $\boldsymbol v=K\boldsymbol w\in BC_{\widetilde\nu}$ such that $\langle\alpha^-,\boldsymbol v(0)\rangle=0$.

In addition, K is a bounded linear operator on $BC_{\widetilde\nu}$.

Proof. Suppose that $\boldsymbol v\in BC_{\widetilde\nu}$ verifies (2.32). Denote $\boldsymbol v_0=\boldsymbol v(0)$. From the Duhamel formula we obtain

$$\boldsymbol v(t)=\mathrm{e}^{tA}\boldsymbol v_0+\int_0^t\mathrm{e}^{(t-\tau)A}\boldsymbol w(\tau)\,\mathrm{d}\tau\ \Rightarrow\ \mathrm{e}^{-tA}\pi_{cu}\boldsymbol v(t)=\pi_{cu}\boldsymbol v_0+\int_0^t\mathrm{e}^{-\tau A}\pi_{cu}\boldsymbol w(\tau)\,\mathrm{d}\tau.\tag{2.33}$$

By assumption, $\|\boldsymbol v(t)\|_{Y^{k+1}\times Y^k}\lesssim\mathrm{e}^{-\widetilde\nu}$, hence from (2.23) we infer $\mathrm{e}^{-tA}\pi_{cu}\boldsymbol v(t)\lesssim(1+t^\kappa)\mathrm{e}^{-\widetilde\nu}$, $\kappa:=\frac12k(k+1)+1$. Passing to the limit $t\to+\infty$ yields

$$\pi_{cu}\boldsymbol v_0=-\int_0^{+\infty}\mathrm{e}^{-\tau A}\pi_{cu}\boldsymbol w(\tau)\,\mathrm{d}\tau.$$

If we require $\langle\alpha^-,\boldsymbol v_0\rangle=0$, this determines uniquely $\boldsymbol v_0=\pi_{cu}\boldsymbol v_0$, hence, using (2.33),

$$\boldsymbol v(t)=K\boldsymbol w(t)=-\int_t^{+\infty}\mathrm{e}^{(t-\tau)A}\pi_{cu}\boldsymbol w(\tau)\,\mathrm{d}\tau+\int_0^t\mathrm{e}^{(t-\tau)A}\pi_s\boldsymbol w(\tau)\,\mathrm{d}\tau.$$

From (2.22) and (2.23) we obtain

$$\|K\boldsymbol w(t)\|_{Y^{k+1}\times Y^k}\lesssim\|\boldsymbol w\|_{BC_{\widetilde\nu}}\cdot\Big(\int_t^{+\infty}(1+(\tau-t)^\kappa)\mathrm{e}^{-\widetilde\nu\tau}\,\mathrm{d}\tau+\int_0^t\mathrm{e}^{(\tau-t)\nu}\mathrm{e}^{-\widetilde\nu\tau}\,\mathrm{d}\tau\Big)$$
$$\lesssim\|\boldsymbol w\|_{BC_{\widetilde\nu}}\cdot\mathrm{e}^{-\widetilde\nu t},$$

so K is a bounded operator. □

Remark 2.8. By linearity the unique solution of (2.32) such that $\langle\alpha^-,\boldsymbol v(0))\rangle=a$ is $\boldsymbol v(t)=(K\boldsymbol w)(t)+\mathrm{e}^{-\nu t}a\mathcal Y^-$.

2.2 Construction of $\boldsymbol{U}_\lambda^a$

As noted earlier, the functions $\boldsymbol{U}_\lambda^a$ were constructed in [27, Section 6]. However, the construction given there does not give the additional regularity or decay, which is required in the present paper. For this reason, we provide here a different construction, which is an adaptation of a classical ODE proof, see for instance [11, Chapter 3.6].

We denote

$$\mathcal{R}(v) := f(W+v) - f(W) - f'(W)v.$$

Lemma 2.9. *Let $\widetilde{\nu} \in (0,\nu)$. There exist $\eta > 0$ such that for every $b \in \mathbb{R}, |b| < \eta$ there is a unique solution $\boldsymbol{v} = \boldsymbol{v}^b \in BC_{\widetilde{\nu}}$ of the equation*

$$\partial_t \boldsymbol{v}(t) = A\boldsymbol{v}(t) + \mathcal{R}(\boldsymbol{v}(t)) \tag{2.34}$$

such that $\langle \alpha^-, \boldsymbol{v}(0)\rangle = b$ and $\|\boldsymbol{v}\|_{BC_{\widetilde{\nu}}} < \eta$. Moreover, $\boldsymbol{v}^b$ is analytic with respect to b.

Proof. Let $T : BC_{\widetilde{\nu}} \times \mathbb{R} \to BC_{\widetilde{\nu}}$ be defined by the formula

$$T(\boldsymbol{v}, b) := \mathrm{e}^{-\nu t} b \mathcal{Y}^- + K(\mathcal{R}(\boldsymbol{v})),$$

where K is the operator from Lemma 2.7. Then $\boldsymbol{v}$ is a solution of (2.34) if and only if $T(\boldsymbol{v}, b) = \boldsymbol{v}$ (see Remark 2.8).

It follows from Lemma A.3 that on some neighbourhood of the origin T is analytic and a uniform contraction with respect to $\boldsymbol{v}$, hence the conclusion follows from the Uniform Contraction Principle, cf. [11, Theorem 2.2]. $\square$

Proposition 2.10. *For any $k \in \mathbb{N}$ there exists $\eta > 0$ and an analytic function*

$$(-\eta,\eta) \ni a \mapsto \boldsymbol{U}^a - \boldsymbol{W} \in Y^{k+1} \times Y^k$$

such that

$$\boldsymbol{U}^0 = \boldsymbol{W}, \tag{2.35}$$

$$\partial_a \boldsymbol{U}^a|_{a=0} = \mathcal{Y}^-, \tag{2.36}$$

$$-\nu a \partial_a \boldsymbol{U}^a = J \circ \mathrm{D}E(\boldsymbol{U}^a). \tag{2.37}$$

Proof. Evaluation at $t = 0$ is a bounded linear operator from $BC_{\widetilde{\nu}}$ to $Y^{k+1} \times Y^k$, hence $\{\boldsymbol{v}^b(0) : b \in (-\eta,\eta)\}$ defines an analytic curve in $Y^{k+1} \times Y^k$. We have $\|\boldsymbol{v}^b\|_{BC_{\widetilde{\nu}}} \lesssim |b|$, so $\|\mathcal{R}(\boldsymbol{v}^b)\|_{BC_{\widetilde{\nu}}} \lesssim |b|^2$. By construction, $\boldsymbol{v}^b$ satisfies the equation

$$\boldsymbol{v}^b = b\mathrm{e}^{-\nu t}\mathcal{Y}^- + K(\mathcal{R}(\boldsymbol{v}^b)),$$

hence $\|\boldsymbol{v}^b - b\mathrm{e}^{-\nu t}\mathcal{Y}^-\|_{BC_{\widetilde{\nu}}} \lesssim |b|^2$, in particular

$$\|\boldsymbol{v}^b(0) - b\mathcal{Y}^-\|_{Y^{k+1}\times Y^k} \lesssim |b|^2.$$

Because of uniqueness in Lemma 2.9, the set $\{\boldsymbol{v}^b(0) : b \in (-\eta,\eta)\}$ is forward invariant if η is small enough, hence for all $b \in (-\eta,\eta)$ there exists a function $b(t)$ such that $\boldsymbol{v}^b(t) = \boldsymbol{v}^{b(t)}(0)$. The value of $b(t)$ is determined by the condition

$$\langle \alpha^-, \boldsymbol{v}^b(t)\rangle = b(t).$$

Differentiating in time this condition we find

$$b'(t) = \frac{\mathrm{d}}{\mathrm{d}t}\langle \alpha^-, \boldsymbol{v}^b(t)\rangle = \big\langle \alpha^-, J \circ \mathrm{D}E(\boldsymbol{W} + \boldsymbol{v}^b(t))\big\rangle = \big\langle \alpha^-, J \circ \mathrm{D}E(\boldsymbol{W} + \boldsymbol{v}^{b(t)}(0))\big\rangle = \psi(b(t)),$$

where ψ is an analytic function, $\psi(0) = 0$ and $\psi'(0) = -\nu$. By Lemma A.1, there exists an analytic change of variable $a = a(b)$ which transforms the equation $b'(t) = \psi(b(t))$ into $a'(t) = -\nu a(t)$ and such that $a(0) = 0$, $a'(0) = 1$. We define

$$\boldsymbol{U}^a := \boldsymbol{W} + \boldsymbol{v}^{b(a)}(0).$$

□

We will denote $\boldsymbol{U}_\lambda^a := (\boldsymbol{U}^a)_\lambda$. Rescaling (2.35), (2.36) and (2.37) we obtain

$$\begin{aligned} \boldsymbol{U}_\lambda^0 &= \boldsymbol{W}_\lambda, \\ \partial_a \boldsymbol{U}_\lambda^a|_{a=0} &= \mathcal{Y}_\lambda^-, \\ \partial_a \boldsymbol{U}_\lambda^a &= -\frac{\lambda}{\nu a} J \circ \mathrm{D}E(\boldsymbol{U}_\lambda^a). \end{aligned} \tag{2.38}$$

Remark 2.11. Note that (2.38) implies that $\boldsymbol{u}(t) = \boldsymbol{U}_\lambda^{\pm\exp(-\frac{\nu}{\lambda}t)}$ is a solution of (NLW) for large t. These are precisely the solutions $\boldsymbol{W}_\lambda^\pm$ mentioned in the Introduction.

2.3 Modulation near the stable manifold

The results of this subsection will not be directly used in the proof of Theorem 1. We include them in the paper for their own interest and because the proofs introduce in a simple setting the main technical ideas required in Section 3.

It is well known since the work of Payne and Sattinger [75] that solutions of energy $< E(\boldsymbol{W})$ leave a neighbourhood of the family of stationary states. The aim of this subsection is to describe an explicit local mechanism leading to this phenomenon, which is robust enough not to be significantly altered by the presence of the second bubble (as will be the case in Section 3).

Note that nothing specific to the wave equation has been used so far, hence one might expect that all the proofs of Section 2 should apply to similar (not necessarily critical) models in the presence of one instability direction near a stationary state.

Lemma 2.12. *Let $\delta_0 > 0$ be sufficiently small. For any $0 \le \delta \le \delta_0$ there exists $0 \le \eta = \eta(\delta) \xrightarrow[\delta\to 0]{} 0$ such that if $\boldsymbol{u} : (t_1, t_2) \to \mathcal{E}$ is a solution of* (NLW) *satisfying for all $t \in (t_1, t_2)$*

$$\|\boldsymbol{u}(t) - \boldsymbol{W}_{\widetilde{\lambda}(t)}\|_\mathcal{E} \le \delta, \qquad \widetilde{\lambda}(t) > 0,$$

then there exist unique functions $\lambda(t) \in C^1((t_1, t_2), (0, +\infty))$ and $a(t) \in C^1((t_1, t_2), \mathbb{R})$ such that for

$$\boldsymbol{g}(t) := \boldsymbol{u}(t) - \boldsymbol{U}_{\lambda(t)}^{a(t)} \tag{2.39}$$

the following holds for all $t \in (t_1, t_2)$:

$$\langle \mathcal{Z}_{\underline{\lambda(t)}}, g(t)\rangle = \langle \alpha_{\lambda(t)}^-, \boldsymbol{g}(t)\rangle = 0, \tag{2.40}$$

$$\|\boldsymbol{g}(t)\|_\mathcal{E} \le \eta, \tag{2.41}$$

$$|\lambda(t)/\widetilde{\lambda}(t) - 1| + |a(t)| \le \eta. \tag{2.42}$$

In addition,

$$|\lambda'(t)| \lesssim \|\boldsymbol{g}(t)\|_\mathcal{E}, \tag{2.43}$$

$$\Big|a'(t) + \frac{\nu}{\lambda(t)} a(t)\Big| \lesssim \frac{1}{\lambda(t)}\big(|a(t)| \cdot \|\boldsymbol{g}(t)\|_\mathcal{E} + \|\boldsymbol{g}(t)\|_\mathcal{E}^2\big). \tag{2.44}$$

Proof. We follow a standard procedure, see for instance [57, Proposition 1].

Step 1. We will first show that for fixed $t_0 \in (t_1, t_2)$ there exist unique $\lambda(t_0)$ and $a(t_0)$ such that (2.40), (2.41) and (2.42) hold for $t = t_0$. Without loss of generality we can assume that $\widetilde{\lambda}(t_0) = 1$ (it suffices to rescale everything).

We consider $\Phi : \mathcal{E} \times \mathbb{R}^2 \to \mathbb{R}^2$ defined as

$$\begin{aligned}\Phi(\boldsymbol{u}_0; l_0, a_0) &= \big(\Phi_1(\boldsymbol{u}_0; l_0, a_0), \Phi_2(\boldsymbol{u}_0; l_0, a_0)\big) \\ &:= \big(\langle \mathrm{e}^{-l_0}\mathcal{Z}_{\underline{\mathrm{e}^{l_0}}}, u_0 - U^{a_0}_{\mathrm{e}^{l_0}}\rangle, \langle \alpha^-_{\mathrm{e}^{l_0}}, \boldsymbol{u}_0 - \boldsymbol{U}^{a_0}_{\mathrm{e}^{l_0}}\rangle\big).\end{aligned}$$

One easily computes:

$$\begin{aligned}\partial_l\Phi_1(\boldsymbol{W}; 0, 0) &= \langle \mathcal{Z}, \Lambda W\rangle > 0, \\ \partial_l\Phi_2(\boldsymbol{W}; 0, 0) &= 0, \\ \partial_a\Phi_1(\boldsymbol{W}; 0, 0) &= 0, \\ \partial_a\Phi_2(\boldsymbol{W}; 0, 0) &= -\langle \alpha^-, \mathcal{Y}^-\rangle = -1.\end{aligned}$$

Applying the Implicit Function Theorem with $\boldsymbol{u}_0 := \boldsymbol{u}(t_0)$ we obtain existence of parameters $a_0 =: a(t_0)$ and $\lambda_0 = \mathrm{e}^{l_0} =: \lambda(t_0)$.

Step 2. We will show that $\lambda(t)$ (equivalently, $l(t) := \log(\lambda(t))$) and $a(t)$ are C^1 functions of t.

Take $t_0 \in (t_1, t_2)$ and let $a_0 := a(t_0)$, $l_0 := \log(\lambda(t_0))$. Define $(\widetilde{l}, \widetilde{a}) : (t_0 - \varepsilon, t_0 + \varepsilon) \to \mathbb{R}^2$ as the solution of the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}(\widetilde{l}(t), \widetilde{a}(t)) = -(\partial_{l,a}\Phi)^{-1}(\mathrm{D}_{\boldsymbol{u}}\Phi)\partial_t\boldsymbol{u}(t)$$

with the initial condition $\widetilde{l}(t_0) = l_0$, $\widetilde{a}(t_0) = a_0$. Notice that $\mathrm{D}_{\boldsymbol{v}}\Phi$ is a continuous functional on $\mathcal{F}$, so we can apply it to $\partial_t\boldsymbol{u}(t)$.

Using the chain rule we get $\frac{\mathrm{d}}{\mathrm{d}t}\Phi(\boldsymbol{u}(t); \widetilde{l}(t), \widetilde{a}(t)) = 0$ for $t \in (t_0 - \varepsilon, t_0 + \varepsilon)$. By continuity, $|\widetilde{l}(t) - l_0| < \eta$ in some neighbourhood of $t = t_0$. Hence, by the uniqueness part of the Implicit Function Theorem, we get $\widetilde{l}(t) = \log\lambda(t)$ and $\widetilde{a}(t) = a(t)$ in some neighbourhood of $t = t_0$. In particular, $\lambda(t)$ and $a(t)$ are of class C^1 in some neighbourhood of t_0.

Step 3. From (2.39) we obtain the following differential equation of the error term $\boldsymbol{g}$:

$$\partial_t\boldsymbol{g} = \partial_t(\boldsymbol{u} - \boldsymbol{U}^a_\lambda) = J \circ \big(\mathrm{D}E(\boldsymbol{u}) - \mathrm{D}E(\boldsymbol{U}^a_\lambda)\big) - \big(\partial_t\boldsymbol{U}^a_\lambda - J \circ \mathrm{D}E(\boldsymbol{U}^a_\lambda)\big).$$

We have

$$\partial_t\boldsymbol{U}^a_\lambda = \lambda'\partial_\lambda\boldsymbol{U}^a_\lambda + a'\partial_a\boldsymbol{U}^a_\lambda = -\lambda' \cdot \frac{1}{\lambda}\Lambda_{\mathcal{E}}\boldsymbol{U}^a_\lambda + a'\partial_a\boldsymbol{U}^a_\lambda, \tag{2.45}$$

so using (2.38) we get

$$\partial_t\boldsymbol{g} = J \circ (\mathrm{D}E(\boldsymbol{U}^a_\lambda + \boldsymbol{g}) - \mathrm{D}E(\boldsymbol{U}^a_\lambda)) + \lambda' \cdot \frac{1}{\lambda}\Lambda_{\mathcal{E}}\boldsymbol{U}^a_\lambda - \Big(a' + \frac{\nu a}{\lambda}\Big)\partial_a\boldsymbol{U}^a_\lambda. \tag{2.46}$$

The first component reads:

$$\partial_t g = \dot g + \lambda'\Lambda U^a_{\underline{\lambda}} + \big(1 + \frac{\lambda a'}{\nu a}\big)\dot U^a_{\underline{\lambda}},$$

hence differentiating in time the first orthogonality relation $\langle \frac{1}{\lambda}\mathcal{Z}_{\underline{\lambda}}, g\rangle = 0$ we obtain

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\langle\frac{1}{\lambda}\mathcal{Z}_{\underline{\lambda}}, g\rangle = -\frac{\lambda'}{\lambda^2}\langle\Lambda_0\mathcal{Z}_{\underline{\lambda}}, g\rangle + \frac{1}{\lambda}\langle\mathcal{Z}_{\underline{\lambda}}, \dot g\rangle + \frac{\lambda'}{\lambda^2}\langle\mathcal{Z}_{\underline{\lambda}}, \Lambda U^a_\lambda\rangle + \Big(\frac{1}{\lambda} + \frac{a'}{\nu a}\Big)\langle\mathcal{Z}_{\underline{\lambda}}, \dot U^a_{\underline{\lambda}}\rangle. \tag{2.47}$$

Differentiating the second orthogonality relation $\langle \alpha_\lambda^-, \boldsymbol{g}\rangle = 0$ and using (2.46) we obtain

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\langle \alpha_\lambda^-, \boldsymbol{g}\rangle = -\frac{\lambda'}{\lambda}\langle \Lambda_{\mathcal{E}^*}\alpha_\lambda^-, \boldsymbol{g}\rangle + \langle \alpha_\lambda^-, J\circ(\mathrm{D}E(\boldsymbol{U}_\lambda^a + \boldsymbol{g}) - \mathrm{D}E(\boldsymbol{U}_\lambda^a))\rangle \\ + \frac{\lambda'}{\lambda}\langle \alpha_\lambda^-, \Lambda_{\mathcal{E}}\boldsymbol{U}_\lambda^a\rangle - \Big(a' + \frac{\nu a}{\lambda}\Big)\langle \alpha_\lambda^-, \partial_a \boldsymbol{U}_\lambda^a\rangle.$$

Together with (2.47) this yields the following linear system for λ' and $\lambda\Big(a' + \frac{\nu a}{\lambda}\Big)$:

$$\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22}\end{pmatrix}\begin{pmatrix} \lambda' \\ \lambda\big(a' + \frac{\nu a}{\lambda}\big)\end{pmatrix} = \begin{pmatrix} -\langle \mathcal{Z}_{\underline{\lambda}}, \dot h\rangle \\ -\lambda\langle \alpha_\lambda^-, J\circ(\mathrm{D}E(\boldsymbol{U}_\lambda^a + \boldsymbol{g}) - \mathrm{D}E(\boldsymbol{U}_\lambda^a))\rangle\end{pmatrix},$$

where

$$\begin{aligned} M_{11} &= \frac{1}{\lambda}\langle \mathcal{Z}_{\underline{\lambda}}, \Lambda U_\lambda^a\rangle - \frac{1}{\lambda}\langle \Lambda_{-1}\mathcal{Z}_{\underline{\lambda}}, g\rangle, \\ M_{12} &= \frac{1}{\nu a}\langle \mathcal{Z}_{\underline{\lambda}}, \dot U_{\underline{\lambda}}^a\rangle, \\ M_{21} &= -\langle \Lambda_{\mathcal{E}^*}\alpha_\lambda^-, \boldsymbol{g}\rangle + \langle \alpha_\lambda^-, \Lambda_{\mathcal{E}}\boldsymbol{U}_\lambda^a\rangle, \\ M_{22} &= -\langle \alpha_\lambda^-, \partial_a \boldsymbol{U}_\lambda^a\rangle. \end{aligned} \tag{2.48}$$

Since $\frac{1}{\lambda}\langle \mathcal{Z}_{\underline{\lambda}}, \Lambda W_\lambda\rangle \gtrsim 1$, $\langle \alpha_\lambda^-, \Lambda_{\mathcal{E}}\boldsymbol{W}\rangle = 0$, $\langle \alpha_\lambda^-, \mathcal{Y}_\lambda^-\rangle = 1$, $\|\Lambda_{\mathcal{E}}\boldsymbol{U}_\lambda^a - \Lambda_{\mathcal{E}}\boldsymbol{W}_\lambda\|_{\mathcal{E}} \lesssim |a|$ and $\|\partial_a \boldsymbol{U}_\lambda^a - \mathcal{Y}_\lambda^-\|_{\mathcal{E}} \lesssim |a|$, we see that

$$\begin{aligned} |M_{11}| &\sim 1, \quad |M_{12}| \lesssim 1, \\ |M_{21}| &\lesssim \|\boldsymbol{g}\|_{\mathcal{E}} + |a|, \quad |M_{22}| \sim 1. \end{aligned}$$

Hence, $\left|\det\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22}\end{pmatrix}\right| \gtrsim 1$ and we obtain

$$\begin{aligned} |\lambda'| &\lesssim |M_{22}|\cdot|\langle \mathcal{Z}_{\underline{\lambda}}, \dot g\rangle| + |M_{12}|\cdot|\lambda\langle \alpha_\lambda^-, J\circ(\mathrm{D}E(\boldsymbol{U}_\lambda^a + \boldsymbol{g}) - \mathrm{D}E(\boldsymbol{U}_\lambda^a))\rangle|, \\ \Big|a + \frac{\lambda a'}{\nu}\Big| &\lesssim |M_{21}|\cdot|\langle \mathcal{Z}_{\underline{\lambda}}, \dot g\rangle| + |M_{11}|\cdot|\lambda\langle \alpha_\lambda^-, J\circ(\mathrm{D}E(\boldsymbol{U}_\lambda^a + \boldsymbol{g}) - \mathrm{D}E(\boldsymbol{U}_\lambda^a))\rangle|. \end{aligned} \tag{2.49}$$

Since $\langle \alpha_\lambda^-, J\circ \mathrm{D}^2E(\boldsymbol{W}_\lambda)\boldsymbol{g}\rangle = -\frac{\nu}{\lambda}\langle \alpha_\lambda^-, \boldsymbol{g}\rangle = 0$, Lemma A.4 implies that

$$|\langle \alpha_\lambda^-, J\circ(\mathrm{D}E(\boldsymbol{U}_\lambda^a + \boldsymbol{g}) - \mathrm{D}E(\boldsymbol{U}_\lambda^a))\rangle| \lesssim \frac{1}{\lambda}\|\boldsymbol{g}\|_{\mathcal{E}}\cdot(|a| + \|\boldsymbol{g}\|_{\mathcal{E}}). \tag{2.50}$$

Now (2.43) and (2.44) follow from (2.48), (2.49) and (2.50).

□

In the rest of this section $\lambda(t)$ and $a(t)$ denote the modulation parameters obtained in Lemma 2.12 and $\boldsymbol{g}(t)$ is the function defined by (2.39).

For given modulation parameters λ and a we define:

$$\beta_\lambda^a := -\frac{\nu}{2\lambda}J\partial_a \boldsymbol{U}_\lambda^a. \tag{2.51}$$

We see that $\beta_\lambda^0 = -\frac{\nu}{2\lambda}J\mathcal{Y}_\lambda^- = \alpha_\lambda^+$, and indeed it turns out that β_λ^a is a refined version of α_λ^+, adapted to the situation when $|a| \gg \|\boldsymbol{h}\|_{\mathcal{E}}$.

Proposition 2.13. *The function*

$$b(t) := \langle \beta_{\lambda(t)}^{a(t)}, \boldsymbol{g}(t)\rangle$$

satisfies

$$\Big|\frac{\mathrm{d}}{\mathrm{d}t}b(t) - \frac{\nu}{\lambda(t)}b(t)\Big| \lesssim \frac{1}{\lambda(t)}\cdot\|\boldsymbol{g}(t)\|_{\mathcal{E}}^2.$$

Proof.

Step 1. We check that

$$|\langle \beta_\lambda^a - \alpha_\lambda^+, \boldsymbol{g}\rangle| \lesssim |a| \cdot \|\boldsymbol{g}\|_{\mathcal{E}}, \tag{2.52}$$

$$|\langle \beta_\lambda^a - \alpha_\lambda^+, \partial_t \boldsymbol{g}\rangle| \lesssim \frac{1}{\lambda}|a| \cdot \|\boldsymbol{g}\|_{\mathcal{E}}, \tag{2.53}$$

$$|\langle \partial_a \beta_\lambda^a, \boldsymbol{g}\rangle| + |\langle \lambda\partial_\lambda \beta_\lambda^a, \boldsymbol{g}\rangle| \lesssim \|\boldsymbol{g}\|_{\mathcal{E}}. \tag{2.54}$$

From Proposition 2.10 we have $\|\beta_1^a - \alpha^+\|_{Y^k\times Y^{k+1}} \lesssim |a|$, and (2.52) follows by rescaling. Analogously one gets (2.54).

Similarly one obtains

$$|\langle \beta_\lambda^a - \alpha_\lambda^+, J\circ(\mathrm{D}E(\boldsymbol{U}_\lambda^a + \boldsymbol{g}) - \mathrm{D}E(\boldsymbol{U}_\lambda^a))\rangle| \lesssim \frac{1}{\lambda}|a| \cdot \|\boldsymbol{g}\|_{\mathcal{E}}, \tag{2.55}$$

$$|\langle \beta_\lambda^a - \alpha_\lambda^+, \Lambda_{\mathcal{E}} \boldsymbol{U}_\lambda^a\rangle| + |\langle \beta_\lambda^a - \alpha_\lambda^+, \partial_a \boldsymbol{U}_\lambda^a\rangle| \lesssim |a|, \tag{2.56}$$

hence (2.53) follows from (2.43) and (2.44).

Note that (2.52) implies in particular that $|\langle \beta_\lambda^a, \boldsymbol{g}\rangle| \lesssim \|\boldsymbol{g}\|_{\mathcal{E}}$ with a universal constant.

Estimates (2.54) follow from the fact that $\|\partial_a\beta_\lambda^a\|_{\mathcal{E}^*} + \|\lambda\partial_\lambda\beta_\lambda^a\|_{\mathcal{E}^*} \lesssim 1$.

Step 2. Consider the case

$$|a(t)| \leq \|\boldsymbol{g}(t)\|_{\mathcal{E}}. \tag{2.57}$$

We have

$$\frac{\mathrm{d}}{\mathrm{d}t}b(t) = \langle \beta_{\lambda(t)}^{a(t)}, \partial_t \boldsymbol{g}(t)\rangle + \lambda'(t)\langle \partial_\lambda \beta_{\lambda(t)}^{a(t)}, \boldsymbol{g}(t)\rangle + a'(t)\langle \partial_a \beta_{\lambda(t)}^{a(t)}, \boldsymbol{g}\rangle. \tag{2.58}$$

From Lemma 2.12 we know that $|\lambda'| \lesssim \|\boldsymbol{g}\|_{\mathcal{E}}$ and $|a'| \lesssim \frac{1}{\lambda}\|\boldsymbol{g}\|_{\mathcal{E}}$. Hence from (2.54) it follows that the last two terms of (2.58) are negligible.

Using (2.57), (2.52) and (2.53) we see that it is sufficient to show that

$$\big|\langle \alpha_\lambda^+, \partial_t \boldsymbol{g}\rangle - \frac{\nu}{\lambda}\langle \alpha_\lambda^+, \boldsymbol{g}\rangle\big| = |\langle \alpha_\lambda^+, \partial_t \boldsymbol{g} - J\circ \mathrm{D}^2E(\boldsymbol{W}_\lambda)\boldsymbol{g}\rangle| \lesssim \frac{1}{\lambda}\|\boldsymbol{g}\|^2. \tag{2.59}$$

This follows easily from (2.46). Indeed, from Lemma A.4 we deduce that

$$|\langle \alpha_\lambda^+, J\circ(\mathrm{D}E(\boldsymbol{U}_\lambda^a + \boldsymbol{g}) - \mathrm{D}E(\boldsymbol{U}_\lambda^a) - \mathrm{D}^2E(\boldsymbol{W}_\lambda)\boldsymbol{g})\rangle| \lesssim \frac{1}{\lambda}\|\boldsymbol{g}\|_{\mathcal{E}}^2.$$

To see that the contribution of the last two terms in (2.46) is negligible, it suffices to use (2.43), (2.44), $|\langle \alpha_\lambda^+, \Lambda_{\mathcal{E}}\boldsymbol{U}_\lambda^a\rangle| \lesssim |a|$ and $|\langle \alpha_\lambda^+, \partial_a \boldsymbol{U}_\lambda^a\rangle| \lesssim 1$.

Step 3. Now consider the case

$$\|\boldsymbol{g}(t)\|_{\mathcal{E}} \leq |a(t)|. \tag{2.60}$$

We can assume that $a \neq 0$ (otherwise $\boldsymbol{u}(t) \equiv \boldsymbol{W}_\lambda$ and the conclusion is obvious).

Using Proposition 2.10 we get

$$\beta_\lambda^a = -\frac{1}{2a}\mathrm{D}E(\boldsymbol{U}_\lambda^a) \quad \Rightarrow \quad b(t) = -\frac{1}{2a(t)}\cdot\langle \mathrm{D}E(\boldsymbol{U}_{\lambda(t)}^{a(t)}), \boldsymbol{g}(t)\rangle.$$

The idea of the proof is that the first factor grows exponentially, while the second does not change much. From (2.44) and (2.60) we obtain $\big|\frac{a'(t)}{a(t)} + \frac{\nu}{\lambda(t)}\big| \lesssim \frac{1}{\lambda(t)}\|\boldsymbol{g}(t)\|$, hence

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}b(t) &= -\frac{a'(t)}{a(t)}b(t) - \frac{1}{2a(t)}\frac{\mathrm{d}}{\mathrm{d}t}\big\langle \mathrm{D}E(\boldsymbol{U}_{\lambda(t)}^{a(t)}), \boldsymbol{g}(t)\big\rangle \\ &= \frac{\nu}{\lambda(t)}b(t) - \frac{1}{2a(t)}\frac{\mathrm{d}}{\mathrm{d}t}\big\langle \mathrm{D}E(\boldsymbol{U}_{\lambda(t)}^{a(t)}), \boldsymbol{g}(t)\big\rangle + \frac{1}{\lambda(t)}O(\|\boldsymbol{g}(t)\|_{\mathcal{E}}^2).\end{aligned}$$

We compute the second term using (2.46):

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathrm{D}E(\boldsymbol{U}_\lambda^a), \boldsymbol{g}\rangle = \langle \mathrm{D}^2E(\boldsymbol{U}_\lambda^a)\partial_t \boldsymbol{U}_\lambda^a, \boldsymbol{g}\rangle \\ + \langle \mathrm{D}E(\boldsymbol{U}_\lambda^a), J\circ(\mathrm{D}E(\boldsymbol{U}_\lambda^a+\boldsymbol{g}) - \mathrm{D}E(\boldsymbol{U}_\lambda^a)) + \lambda'\cdot\frac{1}{\lambda}\Lambda_{\mathcal{E}}\boldsymbol{U}_\lambda^a - \Big(a' + \frac{\nu a}{\lambda}\Big)\partial_a \boldsymbol{U}_\lambda^a\rangle.$$

Observe that

$$\begin{aligned}\langle \mathrm{D}E(\boldsymbol{U}_\lambda^a), \Lambda_{\mathcal{E}}\boldsymbol{U}_\lambda^a\rangle &= -\partial_\lambda E(\boldsymbol{U}_\lambda^a) = 0,\\ \langle \mathrm{D}E(\boldsymbol{U}_\lambda^a), \partial_a\boldsymbol{U}_\lambda^a\rangle &= \partial_a E(\boldsymbol{U}_\lambda^a) = 0.\end{aligned} \tag{2.61}$$

Since $\mathrm{D}E(\boldsymbol{U}_\lambda^a) \in Y^k \times Y^{k+1}$ by Proposition 2.10, Lemma A.4 implies that

$$\big|\big\langle \mathrm{D}E(\boldsymbol{U}_\lambda^a), J\circ\big(\mathrm{D}E(\boldsymbol{U}_\lambda^a+\boldsymbol{g}) - \mathrm{D}E(\boldsymbol{U}_\lambda^a) - \mathrm{D}^2E(\boldsymbol{U}_\lambda^a)\boldsymbol{g}\big)\big\rangle\big| \lesssim \frac{a}{\lambda}\|\boldsymbol{g}\|_{\mathcal{E}}^2,$$

hence using self-adjointness of $\mathrm{D}^2E(\boldsymbol{U}_\lambda^a)$ and anti-self-adjointness of J we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathrm{D}E(\boldsymbol{U}_\lambda^a), \boldsymbol{g}\rangle = \langle \mathrm{D}^2E(\boldsymbol{U}_\lambda^a)\big(\partial_t\boldsymbol{U}_\lambda^a - J\circ \mathrm{D}E(\boldsymbol{U}_\lambda^a)\big), \boldsymbol{g}\rangle + \frac{a}{\lambda}O(\|\boldsymbol{g}\|_{\mathcal{E}}^2).$$

The following estimates hold:

$$\begin{aligned}\|\mathrm{D}^2E(\boldsymbol{U}_\lambda^a)\Lambda_{\mathcal{E}}\boldsymbol{U}_\lambda^a\|_{\mathcal{E}^*} &\lesssim |a|,\\ \|\mathrm{D}^2E(\boldsymbol{U}_\lambda^a)\partial_a\boldsymbol{U}_\lambda^a\|_{\mathcal{E}^*} &\lesssim 1\end{aligned} \tag{2.62}$$

(the first one follows from $\mathrm{D}^2E(\boldsymbol{W}_\lambda)\Lambda_{\mathcal{E}}\boldsymbol{W}_\lambda = 0$). Using (2.45) and (2.62) together with (2.43) and (2.44) concludes the proof. □

As an application of the preceding proposition, we now show that the stable manifold $\boldsymbol{U}_\lambda^a$ is the only source of the lack of coercivity of the energy functional restricted to the trajectories staying close to the family of stationary states.

Given $\boldsymbol{u}_0 \in \mathcal{E}$, let $\boldsymbol{u}(t) : [0, T_+) \to \mathcal{E}$ denote the maximal solution of (NLW) with initial data $\boldsymbol{u}(0) = \boldsymbol{u}_0$. For $\eta > 0$ sufficiently small we define the *centre-stable set* $\mathcal{M}_{\mathrm{cs}}$ as

$$\mathcal{M}_{\mathrm{cs}} := \big\{\boldsymbol{u}_0 : \sup_{0\le t<T_+}\inf_{\lambda>0}\|\boldsymbol{u}(t) - \boldsymbol{W}_\lambda\|_{\mathcal{E}} \le \eta\big\}.$$

Remark 2.14. In the case $N = 3$ it was proved by Krieger, Nakanishi and Schlag [49] that $\mathcal{M}_{\mathrm{cs}}$ is a local C^1 manifold tangent at $\boldsymbol{u}_0 = \boldsymbol{W}$ to $\mathcal{X}_{\mathrm{cs}}$.

Remark 2.15. It is not difficult to see that if $\mathcal{M}_{\mathrm{cs}}$ is a regular hypersurface, then necessarily its tangent space at $\boldsymbol{U}_\lambda^a$ is given by

$$\boldsymbol{U}_\lambda^a + \ker\beta_\lambda^a = \{\boldsymbol{U}_\lambda^a + \boldsymbol{g} : \langle\beta_\lambda^a, \boldsymbol{g}\rangle = 0\}.$$

Hence $b(t)$ is a natural candidate to measure how a trajectory moves away from $\mathcal{M}_{\mathrm{cs}}$.

Corollary 2.16. *If $\eta > 0$ is small enough, then there exists a constant $C_E > 0$ such that*

$$\boldsymbol{u}_0 \in \mathcal{M}_{\mathrm{cs}} \quad\Rightarrow\quad \inf_{\lambda>0, a\in\mathbb{R}}\|\boldsymbol{u}_0 - \boldsymbol{U}_\lambda^a\|_{\mathcal{E}}^2 \le C_E\big(E(\boldsymbol{u}_0) - E(\boldsymbol{W})\big).$$

Proof.

Step 1 – Coercivity. We will prove that if $\|\boldsymbol{g}\|_{\mathcal{E}}$ is small enough and $\langle \mathcal{Z}_{\underline{\lambda}}, g\rangle = \langle \alpha_\lambda^-, \boldsymbol{g}\rangle = 0$, then

$$E(\boldsymbol{U}_\lambda^a + \boldsymbol{g}) - E(\boldsymbol{W}) + 2a\langle \beta_\lambda^a, \boldsymbol{g}\rangle + 2|\langle \beta_\lambda^a, \boldsymbol{g}\rangle|^2 \sim \|\boldsymbol{g}\|_{\mathcal{E}}^2. \tag{2.63}$$

We have $2a\langle \beta_\lambda^a, \boldsymbol{g}\rangle = \langle \mathrm{D}E(\boldsymbol{U}_\lambda^a), \boldsymbol{g}\rangle$, hence Lemma A.5 implies

$$\begin{aligned} E(\boldsymbol{U}_\lambda^a + \boldsymbol{g}) &= E(\boldsymbol{U}_\lambda^a) + \langle \mathrm{D}E(\boldsymbol{U}_\lambda^a), \boldsymbol{g}\rangle + \frac{1}{2}\langle \mathrm{D}^2 E(\boldsymbol{U}_\lambda^a)\boldsymbol{g}, \boldsymbol{g}\rangle + o(\|\boldsymbol{g}\|_{\mathcal{E}}^2) \\ &= E(\boldsymbol{W}) - 2a\langle \beta_\lambda^a, \boldsymbol{g}\rangle + \frac{1}{2}\langle \mathrm{D}^2 E(\boldsymbol{U}_\lambda^a)\boldsymbol{g}, \boldsymbol{g}\rangle + o(\|\boldsymbol{g}\|_{\mathcal{E}}^2). \end{aligned}$$

By (2.52) we have $|\langle \beta_\lambda^a, \boldsymbol{g}\rangle^2 - \langle \alpha_\lambda^+, \boldsymbol{g}\rangle^2| \lesssim |a|\cdot\|\boldsymbol{g}\|^2$, hence Lemma 2.2 yields

$$\frac{1}{2}\langle \mathrm{D}^2 E(\boldsymbol{U}_\lambda^a)\boldsymbol{g}, \boldsymbol{g}\rangle + 2|\langle \beta_\lambda^a, \boldsymbol{g}\rangle|^2 \sim \|\boldsymbol{g}\|_{\mathcal{E}}^2,$$

which implies (2.63).

Step 2 – Differential inequalities. Let $\boldsymbol{g}(t)$, $\lambda(t)$ and $a(t)$ be given by Lemma 2.12. Observe that

$$\int_0^{T_+} \frac{1}{\lambda(t)}\,\mathrm{d}t = +\infty. \tag{2.64}$$

Indeed, if $|\log\lambda(t)|$ is unbounded, then

$$\int_0^{T_+} \frac{1}{\lambda(t)}\,\mathrm{d}t \gtrsim \int_0^{T_+} \frac{\|\boldsymbol{g}(t)\|_{\mathcal{E}}}{\lambda(t)}\,\mathrm{d}t \gtrsim \int_0^{T_+} \frac{|\lambda'(t)|}{\lambda(t)}\,\mathrm{d}t = +\infty.$$

If $|\log\lambda(t)|$ is bounded, then by the Cauchy theory $T_+ = +\infty$ and (2.64) follows.

From Proposition 2.13 it follows that there exists a constant C_1 such that

$$|b(t)| \geq C_1\|\boldsymbol{g}(t)\|_{\mathcal{E}}^2 \quad\Rightarrow\quad \frac{\mathrm{d}}{\mathrm{d}t}|b(t)| \geq \frac{\nu}{2\lambda(t)}|b(t)|, \qquad \forall t \in [0, T_+). \tag{2.65}$$

We will show that there exists a constant C_2 such that

$$|b(t)| \geq C_2\big(E(\boldsymbol{u}_0) - E(\boldsymbol{W})\big) \quad\Rightarrow\quad |b(t)| \geq C_1\|\boldsymbol{g}(t)\|_{\mathcal{E}}^2. \tag{2.66}$$

Indeed, we can rewrite (2.63) as

$$E(\boldsymbol{u}_0) - E(\boldsymbol{W}) + 2a(t)b(t) + 2b(t)^2 \sim \|\boldsymbol{g}\|_{\mathcal{E}}^2, \tag{2.67}$$

hence if $|b(t)| \geq C_2$, then

$$|b(t)|\cdot\big(\frac{1}{C_2} + 2|a(t)| + 2|b(t)|\big) \geq E(\boldsymbol{u}_0) - E(\boldsymbol{W}) + 2a(t)b(t) + 2b(t)^2 \gtrsim \|\boldsymbol{g}\|_{\mathcal{E}}^2,$$

which implies (2.66) since $|a(t)|$ and $|b(t)|$ are small.

Suppose for the sake of contradiction that $b(0) \neq 0$ and $|b(0)| \geq 2C_2\big(E(\boldsymbol{u}_0) - E(\boldsymbol{W})\big)$. Let $t_1 \leq T_+$ be maximal such that

$$b(t) \neq 0,\ |b(t)| \geq C_2\big(E(\boldsymbol{u}_0) - E(\boldsymbol{W})\big), \qquad \forall t \in [0, t_1). \tag{2.68}$$

Of course $t_1 > 0$. Suppose that $t_1 < T_+$. But (2.66) and (2.65) imply that $\frac{\mathrm{d}}{\mathrm{d}t}|b(t)| > 0$ for $t \in [0, t_1]$. In particular, (2.68) cannot break down at $t = t_1$. Thus $t_1 = T_+$ and (2.66) implies that for $t \in [0, T_+)$ there holds $|b(t)| \geq C_1\|\boldsymbol{g}\|_{\mathcal{E}}$. By (2.65) and (2.64), this would imply $|\beta(t)| \xrightarrow[t\to T_+]{} +\infty$, which is absurd.

As a result, $|b(0)| \leq 2C_2\big(E(\boldsymbol{u}_0) - E(\boldsymbol{W})\big)$. Since $|a(0)|$ and $\|\boldsymbol{g}(0)\|_{\mathcal{E}}$ may be assumed as small as we wish, the conclusion follows from (2.67) applied at $t = 0$. $\square$

Remark 2.17. It follows quite easily from Lemma 2.2 that

$$\boldsymbol{g} \in \mathcal{X}_{\mathrm{c}} \Rightarrow \quad \exists b \in \mathbb{R} \ : \ \|\boldsymbol{g} - b\Lambda_{\mathcal{E}}\boldsymbol{W}\|_{\mathcal{E}}^2 \lesssim \frac{1}{2}\langle \mathrm{D}^2 E(\boldsymbol{W})\boldsymbol{g}, \boldsymbol{g}\rangle, \tag{2.69}$$

$$\boldsymbol{g} \in \mathcal{X}_{\mathrm{cs}} \Rightarrow \quad \exists a, b \in \mathbb{R} \ : \ \|\boldsymbol{g} - b\Lambda_{\mathcal{E}}\boldsymbol{W} - a\mathcal{Y}^-\|_{\mathcal{E}}^2 \lesssim \frac{1}{2}\langle \mathrm{D}^2 E(\boldsymbol{W})\boldsymbol{g}, \boldsymbol{g}\rangle. \tag{2.70}$$

Corollary 2.16 provides a nonlinear version of (2.70). By similar methods (analyzing just the *linear* stability and instability components α_λ^+ and α_λ^-) one can prove a nonlinear analogue of (2.69), that is

$$\boldsymbol{u}_0 \in \mathcal{M}_{\mathrm{c}} \quad \Rightarrow \quad \inf_{\lambda > 0} \|\boldsymbol{u}_0 - \boldsymbol{W}_\lambda\|_{\mathcal{E}}^2 \leq C_E\big(E(\boldsymbol{u}_0) - E(\boldsymbol{W})\big),$$

where

$$\mathcal{M}_{\mathrm{c}} := \mathcal{M}_{\mathrm{cs}} \cap \mathcal{M}_{\mathrm{cu}} = \big\{\boldsymbol{u}_0 : \sup_{T_- < t < T_+} \inf_{\lambda > 0} \|\boldsymbol{u}(t) - \boldsymbol{W}_\lambda\|_{\mathcal{E}} \leq \eta\big\}.$$

3 Nonexistence of pure two-bubbles with opposite signs

3.1 Modulation near the sum of two bubbles

Because of a slow decay of W, we will introduce compactly supported approximations of W_λ. Let $R > 0$ be a large constant to be chosen later.

We denote

$$V_R(\lambda_1, \lambda_2)(x) := \begin{cases} W_{\lambda_1}(x) - \zeta(\lambda_1, \lambda_2) & \text{when } |x| \leq R\sqrt{\lambda_1\lambda_2}, \\ 0 & \text{when } |x| \geq R\sqrt{\lambda_1\lambda_2}, \end{cases} \tag{3.1}$$

where

$$\zeta(\lambda_1, \lambda_2) := W_{\lambda_1}(R\sqrt{\lambda_1\lambda_2}) = \frac{1}{\lambda_1^{\frac{N-2}{2}}}\Big(1 + \frac{R^2\lambda_2}{N(N-2)\lambda_1}\Big)^{-\frac{N-2}{2}} = \Big(\lambda_1 + \frac{R^2\lambda_2}{N(N-2)}\Big)^{-\frac{N-2}{2}}.$$

We have $\zeta(\lambda_1, \lambda_2) \sim R^{-(N-2)}\lambda_2^{-\frac{N-2}{2}}$, $\partial_{\lambda_1}\zeta(\lambda_1, \lambda_2) \sim R^{-N}\lambda_2^{-\frac{N}{2}}$ and $\partial_{\lambda_2}\zeta(\lambda_1, \lambda_2) \sim R^{-N}\lambda_2^{-\frac{N}{2}}$.

We will also denote

$$\boldsymbol{V}_R(\lambda_1, \lambda_2) := (V_R(\lambda_1, \lambda_2), 0) \in \mathcal{E}.$$

It is straightforward to check that $V_R(\lambda_1, \lambda_2)$ has weak derivatives $\partial_{\lambda_1}V_R(\lambda_1, \lambda_2)$ and $\partial_{\lambda_2}V_R(\lambda_1, \lambda_2)$, which are given by the formulas:

$$\partial_{\lambda_1}V_R(\lambda_1, \lambda_2)(x) = \begin{cases} -(\Lambda W)_{\underline{\lambda_1}}(x) - \partial_{\lambda_1}\zeta(\lambda_1, \lambda_2) & \text{when } |x| < R\sqrt{\lambda_1\lambda_2}, \\ 0 & \text{when } |x| > R\sqrt{\lambda_1\lambda_2}, \end{cases}$$

$$\partial_{\lambda_2}V_R(\lambda_1, \lambda_2)(x) = \begin{cases} -\partial_{\lambda_2}\zeta(\lambda_1, \lambda_2) & \text{when } |x| < R\sqrt{\lambda_1\lambda_2}, \\ 0 & \text{when } |x| > R\sqrt{\lambda_1\lambda_2}. \end{cases} \tag{3.2}$$

Notice that $\partial_{\lambda_j}V_R(\lambda_1, \lambda_2) \in L^2$ and $\partial_{\lambda_j}V_R(\lambda_1, \lambda_2) \notin \dot H^1$.

In the whole section we will denote $\lambda := \frac{\lambda_1}{\lambda_2}$ and $\mathcal{N}(\boldsymbol{g}, \lambda) := \sqrt{\|\boldsymbol{g}\|_{\mathcal{E}}^2 + \lambda^{\frac{N-2}{2}}}$.

Lemma 3.1. *For $R \gg 1$ and $\lambda \ll 1$ the following estimates are true with constants depending only on the dimension:*

$$\|V_R(\lambda_1, \lambda_2) - W_{\lambda_1}\|_{\dot H^1} \lesssim R^{-\frac{N-2}{2}}\lambda^{\frac{N-2}{4}}, \tag{3.3}$$

$$\|V_R(\lambda_1,\lambda_2) - W_{\lambda_1}\|_{L^\infty} \lesssim R^{-N+2}\lambda_2^{-\frac{N-2}{2}}, \tag{3.4}$$

$$\|\partial_{\lambda_1} V_R(\lambda_1,\lambda_2) + \Lambda W_{\underline{\lambda_1}}\|_{L^\infty(|x|<R\sqrt{\lambda_1\lambda_2})} \lesssim R^{-N}\lambda_2^{-\frac{N}{2}}, \tag{3.5}$$

$$\|V_R(\lambda_1,\lambda_2)\|_{L^1} \lesssim R^2\lambda_2^{\frac{N+2}{2}}\lambda^{\frac{N}{2}}, \tag{3.6}$$

$$\|\partial_{\lambda_1} V_R(\lambda_1,\lambda_2)\|_{L^1} \lesssim R^2\lambda_2^{\frac{N}{2}}\lambda^{\frac{N-2}{2}}. \tag{3.7}$$

Proof. The proof of (3.3), (3.4) and (3.5) is straightforward, see [39, Lemma 2.3]; (3.6) and (3.7) follow from the fact that $|V_R(x)| \lesssim \lambda_1^{-\frac{N-2}{2}} \cdot \big(\frac{|x|}{\lambda_1}\big)^{-N+2}$, $|\partial_{\lambda_1} V_R(x)| \lesssim \lambda_1^{-\frac{N}{2}} \cdot \big(\frac{|x|}{\lambda_1}\big)^{-N+2}$ and $\operatorname{supp}\big(V(x)\big) = \operatorname{supp}\big(\partial_{\lambda_1} V(x)\big) = B(0, R\sqrt{\lambda_1\lambda_2})$. □

We will omit the subscript and write $\boldsymbol{V}(\lambda_1,\lambda_2)$ instead of $\boldsymbol{V}_R(\lambda_1,\lambda_2)$. The approximate solution we will consider is defined as follows:

$$\boldsymbol{U}(\lambda_1,\lambda_2,a_2) := \boldsymbol{U}_{\lambda_2}^{a_2} - \boldsymbol{V}(\lambda_1,\lambda_2).$$

Observe that

$$\partial_{\lambda_1}\boldsymbol{U}(\lambda_1,\lambda_2,a_2) = -\partial_{\lambda_1}\boldsymbol{V}(\lambda_1,\lambda_2), \tag{3.8}$$

$$\partial_{\lambda_2}\boldsymbol{U}(\lambda_1,\lambda_2,a_2) = -\frac{1}{\lambda_2}\Lambda_{\mathcal{E}}\boldsymbol{U}_{\lambda_2}^{a_2} - \partial_{\lambda_2}\boldsymbol{V}(\lambda_1,\lambda_2), \tag{3.9}$$

$$\partial_{a_2}\boldsymbol{U}(\lambda_1,\lambda_2,a_2) = \partial_a\boldsymbol{U}_{\lambda_2}^{a_2} = -\frac{\lambda_2}{\nu a_2} J \circ \mathrm{D}E(\boldsymbol{U}_{\lambda_2}^{a_2}). \tag{3.10}$$

Remark 3.2. The following version of the Implicit Function Theorem has the advantage of providing a lower bound on the size of a ball where it can be applied:

Suppose that X, Y and Z are Banach spaces, $x_0 \in X$, $y_0 \in Y$, $\rho, \eta > 0$ and $\Phi : B(x_0,\rho) \times B(y_0,\eta) \to Z$ is continuous in x and continuously differentiable in y, $\Phi(x_0,y_0) = 0$ and $\mathrm{D}_y\Phi(x_0,y_0) =: L_0$ has a bounded inverse. Suppose that

$$\|L_0 - \mathrm{D}_y\Phi(x,y)\|_Z \leq \frac{1}{3}\|L_0^{-1}\|_{\mathscr{L}(Z,Y)}^{-1} \qquad \text{for } \|x-x_0\|_X < \rho, \|y-y_0\|_Y < \eta, \tag{3.11}$$

$$\|\Phi(x,y_0)\|_Z \leq \frac{\eta}{3}\|L_0^{-1}\|_{\mathscr{L}(Z,Y)}^{-1} \qquad \text{for } \|x-x_0\|_X < \rho. \tag{3.12}$$

Then there exists $y \in C(B(x_0,\rho), B(y_0,\eta))$ such that for $x \in B(x_0,\rho)$, $y = y(x)$ is the unique solution of the equation $\Phi(x,y(x)) = 0$ in $B(y_0,\eta)$. □

The proof is the same as standard proofs of IFT, see for instance [11, Section 2.2].

Lemma 3.3. *Let $\delta_0 > 0$ and $\lambda_0 > 0$ be sufficiently small. For any $0 \leq \delta \leq \delta_0$ and $0 < \widetilde{\lambda} \leq \lambda_0$ there exists $0 \leq \eta = \eta(\delta,\widetilde{\lambda}) \xrightarrow[\delta,\widetilde{\lambda}\to 0]{} 0$ such that if $\boldsymbol{u} : (t_1,t_2) \to \mathcal{E}$ is a solution of* (NLW) *satisfying for all $t \in (t_1,t_2)$*

$$\|\boldsymbol{u}(t) - (-\boldsymbol{W}_{\widetilde{\lambda}_1(t)} + \boldsymbol{W}_{\widetilde{\lambda}_2(t)})\|_{\mathcal{E}} \leq \delta, \qquad 0 < \frac{\widetilde{\lambda}_1(t)}{\widetilde{\lambda}_2(t)} \leq \widetilde{\lambda},$$

then there exist unique functions $\lambda_1(t) \in C^1((t_1,t_2),(0,+\infty))$, $\lambda_2(t) \in C^1((t_1,t_2),(0,+\infty))$ and $a_2(t) \in C^1((t_1,t_2),\mathbb{R})$ such that for

$$\boldsymbol{g}(t) := \boldsymbol{u}(t) - \boldsymbol{U}(\lambda_1,\lambda_2,a_2) \tag{3.13}$$

the following holds for all $t \in (t_1,t_2)$:

$$\langle \mathcal{Z}_{\underline{\lambda_1(t)}}, g(t)\rangle = \langle \mathcal{Z}_{\underline{\lambda_2(t)}}, g(t)\rangle = \langle \alpha_{\lambda_2(t)}^-, \boldsymbol{g}(t)\rangle = 0,$$

$$\|\boldsymbol{g}(t)\|_{\mathcal{E}} \leq \eta,$$
$$|\lambda_1(t)/\widetilde{\lambda}_1(t) - 1| + |\lambda_2(t)/\widetilde{\lambda}_2(t) - 1| + |a_2(t)| \leq \eta.$$

In addition,

$$|\lambda_1'(t)| + |\lambda_2'(t)| \lesssim \mathcal{N}(\boldsymbol{g}(t), \lambda(t)), \tag{3.14}$$
$$\big|a_2'(t) + \frac{\nu}{\lambda_2(t)} a_2(t)\big| \lesssim \frac{1}{\lambda_2(t)}\big(|a_2(t)| \cdot \mathcal{N}(\boldsymbol{g}(t), \lambda(t)) + \mathcal{N}(\boldsymbol{g}(t), \lambda(t))^2\big), \tag{3.15}$$

with constants which may depend on R.

Proof. We will follow the same scheme as in the proof of Lemma 2.12. One additional difficulty is that we cannot reduce by rescaling to modulation near one specific function as we did before.

Step 1. We consider $\Phi : \mathcal{E} \times \mathbb{R}^3 \to \mathbb{R}^3$ defined as

$$\Phi(\boldsymbol{u}_0; l_1, l_2, a_2) = \big(\Phi_1(\boldsymbol{u}_0; l_1, l_2, a_2), \Phi_2(\boldsymbol{u}_0; l_1, l_2, a_2), \Phi_3(\boldsymbol{u}_0; l_1, l_2, a_2)\big)$$
$$:= \big(\langle \frac{1}{\lambda_1}\mathcal{Z}_{\underline{\lambda_1}}, u_0 - U(\lambda_1, \lambda_2, a_2)\rangle, \langle \frac{1}{\lambda_2}\mathcal{Z}_{\underline{\lambda_2}}, u_0 - U(\lambda_1, \lambda_2, a_2)\rangle, \langle \alpha_{\lambda_2}^-, \boldsymbol{u}_0 - \boldsymbol{U}(\lambda_1, \lambda_2, a_2)\rangle\big),$$

where we have already written λ_j instead of e^{l_j} in order to simplify the notation. We will verify that the assumptions (3.11) and (3.12) are satisfied for $x_0 = \boldsymbol{U}(\widetilde{\lambda}_1, \widetilde{\lambda}_2, 0)$, $y_0 = (\widetilde{l}_1, \widetilde{l}_2, 0)$ (where $\widetilde{l}_j := \log \widetilde{\lambda}_j$), ρ small and $\eta = C\rho$ with C a universal constant. We define:

$$M_{11}(\boldsymbol{g}; \lambda_1, \lambda_2, a_2) := \langle \frac{1}{\lambda_1}\mathcal{Z}_{\lambda_1}, \lambda_1\partial_{\lambda_1} V(\lambda_1, \lambda_2)\rangle - \langle \frac{1}{\lambda_1}\Lambda_{-1}\mathcal{Z}_{\underline{\lambda_1}}, g\rangle,$$
$$M_{12}(\boldsymbol{g}; \lambda_1, \lambda_2, a_2) := \langle \frac{1}{\lambda_1}\mathcal{Z}_{\underline{\lambda_1}}, \Lambda U_{\lambda_2}^{a_2} + \lambda_2\partial_{\lambda_2} V(\lambda_1, \lambda_2)\rangle,$$
$$M_{13}(\boldsymbol{g}; \lambda_1, \lambda_2, a_2) := -\langle \frac{1}{\lambda_1}\mathcal{Z}_{\underline{\lambda_1}}, \partial_a U_{\lambda_2}^{a_2}\rangle,$$
$$M_{21}(\boldsymbol{g}; \lambda_1, \lambda_2, a_2) := \langle \frac{1}{\lambda_2}\mathcal{Z}_{\underline{\lambda_2}}, \lambda_1\partial_{\lambda_1} V(\lambda_1, \lambda_2)\rangle,$$
$$M_{22}(\boldsymbol{g}; \lambda_1, \lambda_2, a_2) := \langle \frac{1}{\lambda_2}\mathcal{Z}_{\underline{\lambda_2}}, \Lambda U_{\lambda_2}^{a_2} + \lambda_2\partial_{\lambda_2} V(\lambda_1, \lambda_2)\rangle - \langle \frac{1}{\lambda_2}\Lambda_{-1}\mathcal{Z}_{\underline{\lambda_2}}, g\rangle,$$
$$M_{23}(\boldsymbol{g}; \lambda_1, \lambda_2, a_2) := -\langle \frac{1}{\lambda_2}\mathcal{Z}_{\underline{\lambda_2}}, \partial_a U_{\lambda_2}^{a_2}\rangle,$$
$$M_{31}(\boldsymbol{g}; \lambda_1, \lambda_2, a_2) := \langle \alpha_{\lambda_2}, \lambda_1\partial_{\lambda_1} V(\lambda_1, \lambda_2)\rangle,$$
$$M_{32}(\boldsymbol{g}; \lambda_1, \lambda_2, a_2) := -\langle \Lambda_{\mathcal{E}^*}\alpha_{\lambda_2}, \boldsymbol{g}\rangle + \langle \alpha_{\lambda_2}, \Lambda_{\mathcal{E}}\boldsymbol{U}_{\lambda_2}^{a_2} + \lambda_2\partial_{\lambda_2}\boldsymbol{V}(\lambda_1, \lambda_2)\rangle,$$
$$M_{33}(\boldsymbol{g}; \lambda_1, \lambda_2, a_2) := -\langle \alpha_{\lambda_2}, \partial_a \boldsymbol{U}_{\lambda_2}^{a_2}\rangle,$$

A straightforward computation yields

$$\begin{aligned} |M_{11}| &\sim 1, & |M_{12}| &\lesssim 1, & |M_{13}| &\lesssim 1, \\ |M_{21}| &\lesssim \lambda^{\frac{N}{2}}, & |M_{22}| &\lesssim 1, & |M_{23}| &\lesssim 1, \\ |M_{31}| &\lesssim \lambda^{\frac{N}{2}}, & |M_{32}| &\lesssim \mathcal{N}(\boldsymbol{g}, \lambda) + |a_2|, & |M_{33}| &\sim 1. \end{aligned} \tag{3.16}$$

Using (3.8), (3.9), (3.10) and the fact that $\partial_{l_j} = \lambda_j\partial_{\lambda_j}$ we see that

$$M_{jk}(\boldsymbol{u}_0 - \boldsymbol{U}(\lambda_1, \lambda_2, a_2); \lambda_1, \lambda_2, a_2) = \partial_{l_k}\Phi_j(\boldsymbol{u}_0; l_1, l_2, a_2), \qquad j \in \{1, 2, 3\},\ k \in \{1, 2\},$$
$$M_{j3}(\boldsymbol{u}_0 - \boldsymbol{U}(\lambda_1, \lambda_2, a_2); \lambda_1, \lambda_2, a_2) = \partial_{a_2}\Phi_j(\boldsymbol{u}_0; l_1, l_2, a_2), \qquad j \in \{1, 2, 3\},$$

hence (3.16) implies that the jacobian matrix of Φ with respect to the modulation parameters is uniformly non-degenerate in a neighbourhood of $\boldsymbol{U}(\lambda_1, \lambda_2, a_2)$. This yields parameters $\lambda_1(t_0)$, $\lambda_2(t_0)$ and $a_2(t_0)$, see Remark 3.2.

Step 2. The argument from the proof of Lemma 2.12 shows that $\lambda_1(t)$, $\lambda_2(t)$ and $a_2(t)$ are C^1 functions of $t \in (t_1, t_2)$.

Step 3. From (3.13) we obtain the following differential equation of the error term $\boldsymbol{g}$:

$$\partial_t \boldsymbol{g} = \partial_t(\boldsymbol{u} - \boldsymbol{U}(\lambda_1, \lambda_2, a_2)) = J \circ \mathrm{D}E(\boldsymbol{U}(\lambda_1, \lambda_2, a_2)) - \partial_t \boldsymbol{U}_{\lambda_2}^{a_2}.$$

Using (3.8), (3.9) and (3.10) this can rewritten as

$$\begin{aligned} \partial_t \boldsymbol{g} &= J \circ (\mathrm{D}E(\boldsymbol{U}(\lambda_1, \lambda_2, a_2) + \boldsymbol{g}) - \mathrm{D}E(\boldsymbol{U}_{\lambda_2}^{a_2})) \\ &\quad + \lambda_1' \partial_{\lambda_1} \boldsymbol{V}(\lambda_1, \lambda_2) + \lambda_2' \cdot \Big(\frac{1}{\lambda_2} \Lambda_{\mathcal{E}} \boldsymbol{U}_{\lambda_2}^{a_2} + \partial_{\lambda_2} \boldsymbol{V}(\lambda_1, \lambda_2)\Big) - \big(a_2' + \frac{\nu}{\lambda_2} a_2\big) \partial_a \boldsymbol{U}_{\lambda_2}^{a_2}. \end{aligned} \tag{3.17}$$

The first component reads:

$$\partial_t g = \dot{g} + \lambda_1' \partial_{\lambda_1} V(\lambda_1, \lambda_2) + \lambda_2' \big(\Lambda U_{\underline{\lambda_2}}^{a_2} + \partial_{\lambda_2} V(\lambda_1, \lambda_2)\big) - \big(a_2' + \frac{\nu}{\lambda_2} a_2\big) \partial_a U_{\lambda_2}^{a_2},$$

hence differentiating in time the first orthogonality relation $\langle \frac{1}{\lambda_1} \mathcal{Z}_{\underline{\lambda_1}}, g\rangle = 0$ we obtain

$$\begin{aligned} 0 = \frac{\mathrm{d}}{\mathrm{d}t} \langle \mathcal{Z}_{\underline{\lambda_1}}, g\rangle &= -\frac{\lambda_1'}{\lambda_1^2} \langle \Lambda_{-1} \mathcal{Z}_{\underline{\lambda_1}}, g\rangle + \frac{1}{\lambda_1} \langle \mathcal{Z}_{\underline{\lambda_1}}, \dot{g}\rangle + \frac{\lambda_1'}{\lambda_1} \langle \mathcal{Z}_{\underline{\lambda_1}}, \partial_{\lambda_1} V(\lambda_1, \lambda_2)\rangle \\ &\quad + \frac{\lambda_2'}{\lambda_1} \langle \mathcal{Z}_{\underline{\lambda_1}}, \Lambda U_{\underline{\lambda_2}}^{a_2} + \partial_{\lambda_2} V(\lambda_1, \lambda_2)\rangle - \frac{1}{\lambda_1} \Big(a_2' + \frac{\nu}{\lambda_2} a_2\Big) \langle \mathcal{Z}_{\underline{\lambda_1}}, \partial_a U_{\lambda_2}^{a_2}\rangle, \end{aligned}$$

which can also be written as

$$M_{11} \cdot \lambda_1' + \lambda M_{12} \cdot \lambda_2' + \lambda M_{13} \cdot \lambda_2 \Big(a_2' + \frac{\nu}{\lambda_2} a_2\Big) = -\langle \mathcal{Z}_{\underline{\lambda_1}}, \dot{g}\rangle, \tag{3.18}$$

where for simplicity we write M_{jk} instead of $M_{jk}(\boldsymbol{g}; \lambda_1, \lambda_2, a_2)$. Similarly, differentiating the second orthogonality relation $\langle \frac{1}{\lambda_2} \mathcal{Z}_{\underline{\lambda_2}}, g\rangle = 0$ we obtain

$$\begin{aligned} 0 = \frac{\mathrm{d}}{\mathrm{d}t} \langle \mathcal{Z}_{\underline{\lambda_2}}, g\rangle &= -\frac{\lambda_2'}{\lambda_2^2} \langle \Lambda_{-1} \mathcal{Z}_{\underline{\lambda_2}}, g\rangle + \frac{1}{\lambda_2} \langle \mathcal{Z}_{\underline{\lambda_2}}, \dot{g}\rangle + \frac{\lambda_1'}{\lambda_2} \langle \mathcal{Z}_{\underline{\lambda_2}}, \partial_{\lambda_1} V(\lambda_1, \lambda_2)\rangle \\ &\quad + \frac{\lambda_2'}{\lambda_2} \langle \mathcal{Z}_{\underline{\lambda_2}}, \Lambda U_{\underline{\lambda_2}}^{a_2} + \partial_{\lambda_2} V(\lambda_1, \lambda_2)\rangle - \frac{1}{\lambda_2} \Big(a_2' + \frac{\nu}{\lambda_2} a_2\Big) \langle \mathcal{Z}_{\underline{\lambda_2}}, \partial_a U_{\lambda_2}^{a_2}\rangle, \end{aligned}$$

which can also be written as

$$\frac{1}{\lambda} M_{21} \cdot \lambda_1' + M_{22} \cdot \lambda_2' + M_{23} \cdot \lambda_2 \Big(a_2' + \frac{\nu}{\lambda_2} a_2\Big) = -\langle \mathcal{Z}_{\underline{\lambda_2}}, \dot{g}\rangle. \tag{3.19}$$

Finally, differentiating the third orthogonality relation $\langle \alpha_{\lambda_2}^-, \boldsymbol{g}\rangle = 0$ we obtain

$$\begin{aligned} 0 = \frac{\mathrm{d}}{\mathrm{d}t} \langle \alpha_{\lambda_2}^-, \boldsymbol{g}\rangle &= -\frac{\lambda_2'}{\lambda_2} \langle \Lambda_{\mathcal{E}^*} \alpha_{\lambda_2}^-, \boldsymbol{g}\rangle + \langle \alpha_{\lambda_2}^-, J \circ (\mathrm{D}E(\boldsymbol{U}(\lambda_1, \lambda_2, a_2) + \boldsymbol{g}) - \mathrm{D}E(\boldsymbol{U}_{\lambda_2}^{a_2}))\rangle \\ &+ \lambda_1' \langle \alpha_{\lambda_2}^-, \partial_{\lambda_1} \boldsymbol{V}(\lambda_1, \lambda_2)\rangle + \frac{\lambda_2'}{\lambda_2} \langle \alpha_{\lambda_2}^-, \Lambda_{\mathcal{E}} \boldsymbol{U}_{\lambda_2}^{a_2} + \lambda_2 \partial_{\lambda_2} \boldsymbol{V}(\lambda_1, \lambda_2)\rangle - \Big(a_2' + \frac{\nu a_2}{\lambda_2}\Big) \langle \alpha_{\lambda_2}^-, \partial_a \boldsymbol{U}_{\lambda_2}^{a_2}\rangle, \end{aligned}$$

which can also be written as

$$\frac{1}{\lambda}M_{31}\cdot\lambda_1'+M_{32}\cdot\lambda_2'+M_{33}\cdot\lambda_2\Big(a_2'+\frac{\nu}{\lambda_2}a_2\Big) = -\lambda_2\langle\alpha_{\lambda_2}^-, J\circ(\mathrm{D}E(\boldsymbol{U}(\lambda_1,\lambda_2,a_2)+\boldsymbol{g})-\mathrm{D}E(\boldsymbol{U}_{\lambda_2}^{a_2})\rangle. \tag{3.20}$$

Equations (3.18), (3.19) and (3.20) form a linear system for λ_1', λ_2' and $\lambda_2\Big(a_2'+\frac{\nu}{\lambda_2}a_2\Big)$:

$$\begin{pmatrix} M_{11} & \lambda M_{12} & \lambda M_{13} \\ \frac{1}{\lambda}M_{21} & M_{22} & M_{23} \\ \frac{1}{\lambda}M_{31} & M_{32} & M_{33}\end{pmatrix}\begin{pmatrix}\lambda_1' \\ \lambda_2' \\ a_2+\frac{\lambda_2 a_2'}{\nu}\end{pmatrix} = \begin{pmatrix} -\langle\mathcal{Z}_{\underline{\lambda_1}},\dot g\rangle \\ -\langle\mathcal{Z}_{\underline{\lambda_2}},\dot g\rangle \\ -\lambda_2\langle\alpha_{\lambda_2}^-, J\circ(\mathrm{D}E(\boldsymbol{U}(\lambda_1,\lambda_2,a_2)+\boldsymbol{g})-\mathrm{D}E(\boldsymbol{U}_{\lambda_2}^{a_2}))\rangle\end{pmatrix}.$$

We will check that

$$|\langle\alpha_{\lambda_2}^-, J\circ(\mathrm{D}E(\boldsymbol{U}(\lambda_1,\lambda_2,a_2)+\boldsymbol{g})-\mathrm{D}E(\boldsymbol{U}_{\lambda_2}^{a_2}))\rangle| \lesssim \frac{1}{\lambda_2}\mathcal{N}(\boldsymbol{g},\lambda)(|a_2|+\mathcal{N}(\boldsymbol{g},\lambda)). \tag{3.21}$$

By (2.50), it suffices to show that

$$|\langle\alpha_{\lambda_2}^-, J\circ(\mathrm{D}E(\boldsymbol{U}(\lambda_1,\lambda_2,a_2)+\boldsymbol{g})-\mathrm{D}E(\boldsymbol{U}_{\lambda_2}^{a_2}+\boldsymbol{g}))\rangle| \lesssim \frac{1}{\lambda_2}\mathcal{N}(\boldsymbol{g},\lambda)^2. \tag{3.22}$$

Without loss of generality we can assume that $\lambda_2 = 1$ and $\lambda_1 = \lambda$, hence (3.22) is equivalent to

$$|\langle\mathcal{Y}, -\Delta V(\lambda,1)+f(-V(\lambda,1)+U^{a_2}+g)-f(U^{a_2}+g)\rangle| \lesssim \mathcal{N}(\boldsymbol{g},\lambda)^2. \tag{3.23}$$

We have

$$|\langle\mathcal{Y},\Delta V(\lambda,1)\rangle| = |\langle\Delta\mathcal{Y}, V(\lambda,1)\rangle| \lesssim \lambda^{\frac{N-2}{2}}$$

because of (3.7). For the other term we use the bound

$$|f(-V(\lambda,1)+U^{a_2}+g)-f(U^{a_2}+g)| \lesssim (f'(U^{a_2})+f'(g))V(\lambda,1)+f(V(\lambda,1)).$$

From (3.6) we obtain $|\langle\mathcal{Y}, f'(U^{a_2})V(\lambda,1)\rangle| \lesssim \|V(\lambda,1)\|_{L^1} \lesssim \mathcal{N}(\boldsymbol{g},\lambda)^2$. Using Hölder we compute

$$|\langle\mathcal{Y}, f'(g)\cdot V(\lambda,1)\rangle| \lesssim \|f'(g)\|_{L^{\frac{N}{2}}}\cdot\|V(\lambda,1)\|_{L^{\frac{N}{N-2}}} \lesssim \|\boldsymbol{g}\|_{\mathcal{E}}^{\frac{4}{N-2}}\cdot\lambda^{\frac{N-2}{2}}|\log\lambda| \lesssim \mathcal{N}(\boldsymbol{g},\lambda)^2.$$

Finally, $|\langle\mathcal{Y}, f(V(\lambda,1))\rangle| \lesssim \|f(W_\lambda)\|_{L^1} \lesssim \lambda^{\frac{N-2}{2}}$. This finishes the proof of (3.23), hence we have shown (3.21).

Consider the inverse matrix

$$\begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33}\end{pmatrix} := \begin{pmatrix} M_{11} & \lambda M_{12} & \lambda M_{13} \\ \frac{1}{\lambda}M_{21} & M_{22} & M_{23} \\ \frac{1}{\lambda}M_{31} & M_{32} & M_{33}\end{pmatrix}^{-1}.$$

From (3.16) we obtain

$$\begin{aligned} |P_{11}| &\lesssim 1, & |P_{12}| &\lesssim 1, & |P_{13}| &\lesssim 1, \\ |P_{21}| &\lesssim 1, & |P_{22}| &\lesssim 1, & |P_{23}| &\lesssim 1, \\ |P_{31}| &\lesssim \mathcal{N}(\boldsymbol{g},\lambda)+|a_2|, & |P_{32}| &\lesssim \mathcal{N}(\boldsymbol{g},\lambda)+|a_2|, & |P_{33}| &\lesssim 1, \end{aligned}$$

hence (3.21) yields (3.14) and (3.15).

□

We finish this subsection by analyzing the stability and instability components at both scales $\lambda_1(t)$ and $\lambda_2(t)$. At the scale $\lambda_2(t)$ we use the refined component $\beta_{\lambda_2}^{a_2}$ introduced in Section 2, see (2.51).

Proposition 3.4. *The functions*

$$a_1^-(t) := \langle \alpha_{\lambda_1(t)}^-, \boldsymbol{g}(t)\rangle, \qquad a_1^+(t) := \langle \alpha_{\lambda_1(t)}^+, \boldsymbol{g}(t)\rangle, \qquad b_2(t) := \langle \beta_{\lambda_2(t)}^{a_2(t)}, \boldsymbol{g}(t)\rangle$$

satisfy

$$\Big|\frac{\mathrm{d}}{\mathrm{d}t}a_1^-(t) + \frac{\nu}{\lambda_1(t)}a_1^-(t)\Big| \lesssim \frac{1}{\lambda_1(t)}\mathcal{N}(\boldsymbol{g}(t), \lambda(t))^2, \tag{3.24}$$

$$\Big|\frac{\mathrm{d}}{\mathrm{d}t}a_1^+(t) - \frac{\nu}{\lambda_1(t)}a_1^+(t)\Big| \lesssim \frac{1}{\lambda_1(t)}\mathcal{N}(\boldsymbol{g}(t), \lambda(t))^2, \tag{3.25}$$

$$\Big|\frac{\mathrm{d}}{\mathrm{d}t}b_2(t) - \frac{\nu}{\lambda_2(t)}b_2(t)\Big| \lesssim \frac{1}{\lambda_2(t)}\mathcal{N}(\boldsymbol{g}(t), \lambda(t))^2, \tag{3.26}$$

with constants eventually depending on R.

Proof.

Step 1. Directly from the definition of $a_1^-(t)$ we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}a_1^-(t) = -\frac{\lambda_1'(t)}{\lambda_1(t)}\langle \Lambda_{\mathcal{E}^*}\alpha_{\lambda_1(t)}^-, \boldsymbol{g}(t)\rangle + \langle \alpha_{\lambda_1(t)}^-, \partial_t \boldsymbol{g}(t)\rangle.$$

The first term is negligible due to (3.14). We compute the second term using (3.17). We begin by treating the terms in the second line of (3.17). Since $|\lambda_1'| + |\lambda_2'| \lesssim 1$ and $\big|a_2' + \frac{\nu}{\lambda_2}a_2\big| \lesssim \frac{1}{\lambda_2}$ (of course Lemma 3.3 provides better estimates, but we do not need it here), it suffices to check that

$$\begin{aligned}&|\langle \alpha_{\lambda_1}^-, \partial_{\lambda_1}\boldsymbol{V}(\lambda_1,\lambda_2)\rangle| + |\langle \alpha_{\lambda_1}^-, \partial_{\lambda_2}\boldsymbol{V}(\lambda_1,\lambda_2)\rangle| \\ &+|\langle \alpha_{\lambda_1}^-, \frac{1}{\lambda_2}\Lambda_{\mathcal{E}}\boldsymbol{U}_{\lambda_2}^{a_2}\rangle| + |\langle \alpha_{\lambda_1}^-, \frac{1}{\lambda_2}\partial_{a_2}\boldsymbol{U}_{\lambda_2}^{a_2}\rangle| \lesssim \frac{1}{\lambda_1}\cdot\big(\frac{\lambda_1}{\lambda_2}\big)^{\frac{N-2}{2}}.\end{aligned}$$

The estimate is invariant by rescaling both λ_1 and λ_2, hence we can assume that $\lambda_2 = 1$ and $\lambda_1 = \lambda$. For the first term we use (3.5) and rapid decay of $\mathcal{Y}$. Estimating the other terms is straightforward.

Now consider the first line of (3.17). It follows from (2.19) that it suffices to show that

$$\big|\big\langle \alpha_{\lambda_1}^-, J\circ\big(\mathrm{D}E(\boldsymbol{U}(\lambda_1,\lambda_2,a_2)+\boldsymbol{g}) - \mathrm{D}E(\boldsymbol{U}_{\lambda_2}^{a_2}) - \mathrm{D}^2E(\boldsymbol{W}_{\lambda_1})\boldsymbol{g}\big)\big\rangle\big| \lesssim \frac{1}{\lambda_1}\mathcal{N}(\boldsymbol{g},\lambda)^2,$$

which is equivalent to

$$|\langle \mathcal{Y}_{\lambda_1}, f(U(\lambda_1,\lambda_2,a_2)+g) - f(U_{\lambda_2}^{a_2}) - \Delta V(\lambda_1,\lambda_2) - f'(W_{\lambda_1})g\rangle| \lesssim \mathcal{N}(\boldsymbol{g},\lambda)^2.$$

We can assume that $\lambda_2 = 1$ and $\lambda_1 = \lambda$. By the triangle inequality, it suffices to check that

$$|\langle \mathcal{Y}_\lambda, \Delta V(\lambda,1) + f(V(\lambda,1))\rangle| \lesssim \mathcal{N}(\boldsymbol{g},\lambda)^2, \tag{3.27}$$

$$|\langle \mathcal{Y}_\lambda, f(U(\lambda,1,a_2)+g) - f(U(\lambda,1,a_2)) - f'(U(\lambda,1,a_2))g\rangle| \lesssim \mathcal{N}(\boldsymbol{g},\lambda)^2, \tag{3.28}$$

$$|\langle \mathcal{Y}_\lambda, f(U(\lambda,1,a_2)) - f(U^{a_2}) + f(V(\lambda,1))\rangle| \lesssim \mathcal{N}(\boldsymbol{g},\lambda)^2, \tag{3.29}$$

$$|\langle \mathcal{Y}_\lambda, \big(f'(U(\lambda,1,a_2))\big) - f'(W_\lambda)\big)g\rangle| \lesssim \mathcal{N}(\boldsymbol{g},\lambda)^2. \tag{3.30}$$

Notice that $|f(W_\lambda) - f(V(\lambda,1))| \lesssim f'(W_\lambda)| \cdot |W_\lambda - V(\lambda)| \lesssim f'(W_\lambda)$, where the last inequality follows from (3.4). Together with the fact that $\Delta(W_\lambda) + f(W_\lambda) = 0$ this implies

$$\begin{aligned}\big|\langle \mathcal{Y}_\lambda, \big(\Delta V(\lambda,1) + f(V(\lambda,1))\big)\rangle\big| &\lesssim \big|\langle \mathcal{Y}_\lambda, \Delta\big(W_\lambda - V(\lambda,1)\big)\rangle\big| + \big|\langle \mathcal{Y}_\lambda, f(W_\lambda) - f(V(\lambda,1))\rangle\big| \\ &\lesssim \|\Delta Y_\lambda\|_{L^1} + \|f'(W_\lambda)\mathcal{Y}_\lambda\|_{L^1} \lesssim \lambda^{\frac{N-2}{2}},\end{aligned}$$

which proves (3.27).

To fix ideas, notice that while proving the remaining inequalities we can restrict our attention to the region $|x| \leq c\sqrt{\lambda}$ where $c > 0$ is a small constant (the region $|x| \geq c\sqrt{\lambda}$ is negligible thanks to the rapid decay of $\mathcal{Y}$). In this region we have $W_\lambda \geq V(\lambda,1) \gtrsim 1$ and $|U(\lambda,1,a_2) + W_\lambda| \leq \frac{1}{2}W_\lambda$ pointwise.

Inequality (3.29) follows immediately from

$$|f(U(\lambda,1,a_2)) - f(U^{a_2}) + f(V(\lambda,1))| = |f(U^{a_2} - V(\lambda,1)) - f(U^{a_2}) + f(V(\lambda,1))| \lesssim f'(W_\lambda).$$

We have the bound

$$|f'(U(\lambda,1,a_2)) - f'(W_\lambda)| \lesssim (|f''(W_\lambda)| + |f''(U(\lambda,1,a_2))|) \cdot |U(\lambda,1,a_2) + W_\lambda| \lesssim |f''(W_\lambda)|$$

(even in the case $N \geq 6$ when f'' is a negative power). Using Hölder and the fact that $\|\mathcal{Y}_\lambda \cdot f''(W_\lambda)\|_{L^{\frac{2N}{N+2}}} \lesssim \lambda^{\frac{N-2}{2}}$, this implies (3.30)

For (3.28), we consider separately the cases $N \in \{3,4,5\}$ and $N \geq 6$. In the first case, (3.28) follows from the pointwise bound

$$|f(U(\lambda,1,a_2) + g) - f(U(\lambda,1,a_2)) - f'(U(\lambda,1,a_2))g| \lesssim |f''(U(\lambda,1,a_2))| \cdot |g|^2 + f(|g|).$$

In the case $N \geq 6$ we still have

$$|f(U(\lambda,1,a_2) + g) - f(U(\lambda,1,a_2)) - f'(U(\lambda,1,a_2))g| \lesssim |f''(U(\lambda,1,a_2))| \cdot |g|^2,$$

even if f'' is a negative power. This yields (3.28).

This finishes the proof of (3.24) and the proof of (3.25) is almost the same.

Step 2. The proof of (3.26) is close to the proof of Proposition 2.13, but there will be more error terms to estimate. First we need to show that

$$|\langle \beta^{a_2}_{\lambda_2} - \alpha^+_{\lambda_2}, \partial_t \boldsymbol{g}\rangle| \lesssim \frac{1}{\lambda_2}|a_2| \cdot \mathcal{N}(\boldsymbol{g},\lambda). \tag{3.31}$$

Since $\|\beta^{a_2}_1 - \alpha^+\|_{L^\infty \times L^\infty} \lesssim |a_2|$, the proof of (3.22) gives

$$|\langle \beta^{a_2}_{\lambda_2} - \alpha^+_{\lambda_2}, J \circ (\mathrm{D}E(\boldsymbol{U}(\lambda_1,\lambda_2,a_2) + \boldsymbol{g}) - \mathrm{D}E(\boldsymbol{U}^{a_2}_{\lambda_2} + \boldsymbol{g}))\rangle| \lesssim \frac{1}{\lambda_2}|a_2| \cdot \mathcal{N}(\boldsymbol{g},\lambda)^2 \ll \frac{1}{\lambda_2}|a_2| \cdot \mathcal{N}(\boldsymbol{g},\lambda).$$

Using (2.55), we obtain

$$|\langle \beta^{a_2}_{\lambda_2} - \alpha^+_{\lambda_2}, J \circ (\mathrm{D}E(\boldsymbol{U}(\lambda_1,\lambda_2,a_2) + \boldsymbol{g}) - \mathrm{D}E(\boldsymbol{U}^{a_2}_{\lambda_2}))\rangle| \lesssim \frac{1}{\lambda_2}|a_2| \cdot \mathcal{N}(\boldsymbol{g},\lambda).$$

Similarly one obtains

$$|\langle \beta^{a_2}_{\lambda_2} - \alpha^+_{\lambda_2}, \partial_{\lambda_1}\boldsymbol{V}(\lambda_1,\lambda_2)\rangle| + |\langle \beta^{a_2}_{\lambda_2} - \alpha^+_{\lambda_2}, \partial_{\lambda_2}\boldsymbol{V}(\lambda_1,\lambda_2)\rangle| \ll \frac{1}{\lambda_2}|a_2|,$$

hence (3.31) follows from (2.56), (3.14) and (3.15).

Step 3. Suppose that

$$|a_2(t)| \leq \mathcal{N}(\boldsymbol{g}(t), \lambda(t)). \tag{3.32}$$

We have

$$\frac{\mathrm{d}}{\mathrm{d}t} b_2(t) = \langle \beta_{\lambda_2(t)}^{a_2(t)}, \partial_t \boldsymbol{g}(t)\rangle + \lambda_2'(t)\langle \partial_\lambda \beta_{\lambda_2(t)}^{a_2(t)}, \boldsymbol{g}(t)\rangle + a_2'(t)\langle \partial_a \beta_{\lambda_2(t)}^{a_2(t)}, \boldsymbol{g}\rangle.$$

From Lemma 3.3 we know that $|\lambda_2'| \lesssim \mathcal{N}(\boldsymbol{g}, \lambda)$ and $|a_2'| \lesssim \frac{1}{\lambda_2}\mathcal{N}(\boldsymbol{g}, \lambda)$. Hence from (2.54) it follows that the last two terms of (3.26) are negligible.

Using (3.32), (2.52) and (3.31) we see that it is sufficient to show that

$$\big|\langle \alpha_{\lambda_2}^+, \partial_t \boldsymbol{g}\rangle - \frac{\nu}{\lambda_2}\langle \alpha_{\lambda_2}^+, \boldsymbol{g}\rangle\big| = |\langle \alpha_{\lambda_2}^+, \partial_t \boldsymbol{g} - J \circ \mathrm{D}^2 E(\boldsymbol{W}_{\lambda_2})\boldsymbol{g}\rangle| \lesssim \frac{1}{\lambda_2}\mathcal{N}(\boldsymbol{g}, \lambda)^2.$$

We develop $\partial_t \boldsymbol{g}$ using (3.17). Consider first the terms in the second line of (3.17). From (3.7) and (3.14) we have

$$|\langle \alpha_{\lambda_2}^+, \lambda_1' \partial_{\lambda_1} \boldsymbol{V}(\lambda_1, \lambda_2)\rangle| \lesssim \frac{1}{\lambda_2}\mathcal{N}(\boldsymbol{g}, \lambda)^2.$$

Since $|\partial_{\lambda_2} V(\lambda_1, \lambda_2)| \lesssim \lambda_2^{-\frac{N}{2}}$, see (3.2), using (3.14) we get

$$|\langle \alpha_{\lambda_2}^+, \lambda_2' \partial_{\lambda_1} \boldsymbol{V}(\lambda_1, \lambda_2)\rangle| \lesssim \frac{1}{\lambda_2}\mathcal{N}(\boldsymbol{g}, \lambda)^2.$$

The other two terms have already appeared in the proof of Proposition 2.13, see (2.59).

Consider now the first line of (3.17). From Lemma A.4 we deduce that

$$|\langle \alpha_{\lambda_2}^+, J \circ (\mathrm{D}E(\boldsymbol{U}_{\lambda_2}^{a_2} + \boldsymbol{g}) - \mathrm{D}E(\boldsymbol{U}_{\lambda_2}^{a_2}) - \mathrm{D}^2 E(\boldsymbol{W}_{\lambda_2})\boldsymbol{g})\rangle| \lesssim \frac{1}{\lambda_2}\mathcal{N}(\boldsymbol{g}, \lambda)^2,$$

hence it suffices to check that

$$|\langle \alpha_{\lambda_2}^+, J \circ (\mathrm{D}E(\boldsymbol{U}(\lambda_1, \lambda_2, a_2) + \boldsymbol{g}) - \mathrm{D}E(\boldsymbol{U}_{\lambda_2}^{a_2} + \boldsymbol{g}))\rangle| \lesssim \frac{1}{\lambda_2}\mathcal{N}(\boldsymbol{g}, \lambda)^2,$$

whose proof is the same as the proof of (3.22).

Step 4. Now we consider the case

$$\mathcal{N}(\boldsymbol{g}(t), \lambda(t)) \leq |a_2(t)|, \tag{3.33}$$

in particular $a_2 \neq 0$.

Recall that (see Proposition 2.10)

$$\beta_{\lambda_2}^{a_2} = -\frac{1}{2a_2}\mathrm{D}E(\boldsymbol{U}_{\lambda_2}^{a_2}) \quad \Rightarrow \quad b_2(t) = -\frac{1}{2a_2(t)} \cdot \langle \mathrm{D}E(\boldsymbol{U}_{\lambda_2(t)}^{a_2(t)}), \boldsymbol{g}(t)\rangle.$$

From (3.15) and (3.33) we obtain $\big|\frac{a_2'(t)}{a_2(t)} + \frac{\nu}{\lambda_2(t)}\big| \lesssim \frac{1}{\lambda_2(t)}\mathcal{N}(g(t), \lambda(t))$, hence

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} b_2(t) &= -\frac{a_2'(t)}{a_2(t)} b_2(t) - \frac{1}{2a_2(t)}\frac{\mathrm{d}}{\mathrm{d}t}\big\langle \mathrm{D}E(\boldsymbol{U}_{\lambda(t)}^{a_2(t)}), \boldsymbol{g}(t)\big\rangle \\ &= \frac{\nu}{\lambda_2(t)} b_2(t) - \frac{1}{2a_2(t)}\frac{\mathrm{d}}{\mathrm{d}t}\big\langle \mathrm{D}E(\boldsymbol{U}_{\lambda_2(t)}^{a_2(t)}), \boldsymbol{g}(t)\big\rangle + \frac{1}{\lambda_2(t)} O(\mathcal{N}(\boldsymbol{g}(t), \lambda(t))^2).\end{aligned}$$

We compute the second term using (3.17) and (2.61):

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathrm{D}E(\boldsymbol{U}_{\lambda_2}^{a_2}), \boldsymbol{g}\rangle &= \langle \mathrm{D}^2 E(\boldsymbol{U}_{\lambda_2}^{a_2})\partial_t \boldsymbol{U}_{\lambda_2}^{a_2}, \boldsymbol{g}\rangle + \langle \mathrm{D}E(\boldsymbol{U}_{\lambda}^{a}), \\ &J\circ(\mathrm{D}E(\boldsymbol{U}(\lambda_1,\lambda_2,a_2)+\boldsymbol{g}) - \mathrm{D}E(\boldsymbol{U}_{\lambda_2}^{a_2})) + \lambda_1'\partial_{\lambda_1}\boldsymbol{V}(\lambda_1,\lambda_2) + \lambda_2'\partial_{\lambda_2}\boldsymbol{V}(\lambda_1,\lambda_2)\rangle.\end{aligned} \tag{3.34}$$

We have to prove that $\left|\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathrm{D}E(\boldsymbol{U}_{\lambda_2}^{a_2}), \boldsymbol{g}\rangle\right| \lesssim \frac{a_2}{\lambda_2}\mathcal{N}(\boldsymbol{g},\lambda)^2$. Until the end of this proof "negligible" means $\lesssim \frac{a_2}{\lambda_2}\mathcal{N}(\boldsymbol{g},\lambda)^2$.

From (3.7) and (3.2) it follows that

$$|\langle \mathrm{D}(\boldsymbol{U}_{\lambda_2}^{a_2}), \partial_{\lambda_1}\boldsymbol{V}(\lambda_1,\lambda_2)\rangle| \lesssim \frac{1}{\lambda_2}\lambda^{\frac{N-2}{2}}, |\langle \mathrm{D}(\boldsymbol{U}_{\lambda_2}^{a_2}), \partial_{\lambda_2}\boldsymbol{V}(\lambda_1,\lambda_2)\rangle| \lesssim \frac{1}{\lambda_2}\lambda^{\frac{N}{2}}.$$

By (3.14) and (3.33), the contribution of the last two terms in (3.34) is negligible.

Next, we will show that

$$|\langle \mathrm{D}(\boldsymbol{U}_{\lambda_2}^{a_2}), J\circ(\mathrm{D}E(\boldsymbol{U}(\lambda_1,\lambda_2,a_2)+\boldsymbol{g}) - \mathrm{D}E(\boldsymbol{U}_{\lambda_2}^{a_2}+\boldsymbol{g}))\rangle| \lesssim \frac{a_2}{\lambda_2}\mathcal{N}(\boldsymbol{g},\lambda)^2.$$

We can assume that $\lambda_2 = 1$ and $\lambda_1 = \lambda$, hence we have to prove that

$$|\langle \dot{U}^{a_2}, f(U(\lambda,1,a_2)+g) - f(U_{\lambda_2}^{a_1}+g)\rangle| \lesssim a_2\mathcal{N}(\boldsymbol{g},\lambda)^2. \tag{3.35}$$

In the region $|x| > R\sqrt{\lambda}$ the integrand equals 0. In the region $|x| \leq R\sqrt{\lambda}$ we have a pointwise bound

$$|f(U(\lambda,1,a_2)+g) - f(U^{a_2}+g)| \lesssim f'(U^{a_2}+g)W_\lambda + f(W_\lambda) \lesssim (f'(U^{a_2}) + f'(g))W_\lambda + f(W_\lambda).$$

Recall that $\|\dot{U}^{a_2}\|_{L^\infty} \lesssim |a_2|$ and $\|U^{a_2}\|_{L^\infty} \lesssim 1$. Thus

$$\begin{aligned}|\langle|\dot{U}^{a_2}|, f'(U^{a_2})W_\lambda\rangle| &\lesssim |a_2|\cdot\|W_\lambda\|_{L^1(|x|\leq R\sqrt{\lambda})} \sim |a_2|\lambda^{\frac{N-2}{2}}, \\ |\langle|\dot{U}^{a_2}|, f'(g)W_\lambda\rangle| &\lesssim |a_2|\cdot\|f'(g)\|_{L^{\frac{N}{2}}}\cdot\|W_\lambda\|_{L^{\frac{N}{N-2}}(|x|\leq R\sqrt{\lambda})} \\ &\lesssim |a_2|\cdot\|g\|_{\dot{H}^1}^{\frac{4}{N-2}}\cdot\lambda^{\frac{N-2}{2}}|\log\lambda| \lesssim |a_2|\mathcal{N}(\boldsymbol{g},\lambda)^2, \\ |\langle|\dot{U}^{a_2}|, f(W_\lambda)\rangle| &\lesssim |a_2|\cdot\|f(W_\lambda)\|_{L^1} \sim |a_2|\lambda^{\frac{N-2}{2}}.\end{aligned}$$

This proves (3.35).

In order to finish the proof, it suffices to check that

$$|\langle \mathrm{D}^2E(\boldsymbol{U}_{\lambda_2}^{a_2})\partial_t\boldsymbol{U}_{\lambda_2}^{a_2}, \boldsymbol{g}\rangle + \langle \mathrm{D}E(\boldsymbol{U}_{\lambda_2}^{a_2}), J\circ(\mathrm{D}E(\boldsymbol{U}_{\lambda_2}^{a_2}+\boldsymbol{g}) - \mathrm{D}E(\boldsymbol{U}_{\lambda_2}^{a_2}))\rangle| \lesssim \frac{a_2}{\lambda_2}\mathcal{N}(\boldsymbol{g},\lambda)^2,$$

which is achieved exactly as in the last part of the proof of Proposition 2.13. □

3.2 Coercivity near the sum of two bubbles

We have the following analogue of Lemma 2.2:

Lemma 3.5. *There exist constants $\lambda_0, \eta > 0$ such that if $\lambda = \frac{\lambda_1}{\lambda_2} < \lambda_0$ and $\|\boldsymbol{U} - (\boldsymbol{W}_{\lambda_2} - \boldsymbol{W}_{\lambda_1})\|_{\mathcal{E}} < \eta$, then for all $\boldsymbol{g}\in\mathcal{E}$ such that $\langle \mathcal{Z}_{\underline{\lambda_1}}, g\rangle = \langle \mathcal{Z}_{\underline{\lambda_2}}, g\rangle = 0$ there holds*

$$\frac{1}{2}\langle \mathrm{D}^2E(\boldsymbol{U})\boldsymbol{g},\boldsymbol{g}\rangle + 2\big(\langle\alpha_{\lambda_1}^-,\boldsymbol{g}\rangle^2 + \langle\alpha_{\lambda_1}^+,\boldsymbol{g}\rangle^2 + \langle\alpha_{\lambda_2}^-,\boldsymbol{g}\rangle^2 + \langle\alpha_{\lambda_2}^+,\boldsymbol{g}\rangle^2\big) \gtrsim \|\boldsymbol{g}\|_{\mathcal{E}}^2.$$

Proof.

Step 1. Without loss of generality we can assume that $\lambda_2 = 1$ and $\lambda_1 = \lambda$. Consider the operator H_λ defined by the following formula:

$$H_\lambda := \begin{pmatrix} -\Delta - f'(W_\lambda) - f'(W) & 0 \\ 0 & \mathrm{Id} \end{pmatrix}.$$

We will show that for any $c > 0$ there holds

$$|\langle \mathrm{D}^2 E(\boldsymbol{U})\boldsymbol{g}, \boldsymbol{g}\rangle - \langle H_\lambda \boldsymbol{g}, \boldsymbol{g}\rangle| \leq c\|\boldsymbol{g}\|_{\mathcal{E}}^2, \qquad \forall \boldsymbol{g} \in \mathcal{E}, \tag{3.36}$$

provided that η and λ_0 are small enough. By Hölder and Sobolev, it suffices (eventually changing c) to check that

$$\|f'(U) - f'(W_\lambda) - f'(W)\|_{L^{\frac{N}{2}}} \leq c.$$

Since (by pointwise estimates)

$$\|f'(U) - f'(W - W_\lambda)\|_{L^{\frac{N}{2}}} \lesssim \max(\eta, f'(\eta)),$$

this will in turn follow from

$$\|f'(W - W_\lambda) - f'(W_\lambda) - f'(W)\|_{L^{\frac{N}{2}}} \leq c. \tag{3.37}$$

We consider separately the regions $|x| \leq \sqrt{\lambda}$ and $|x| \geq \sqrt{\lambda}$. In both cases we will use the fact that

$$\begin{aligned} |l| \lesssim |k| \quad &\Rightarrow \quad \big|f'(k+l) - f'(k) - f'(l)\big| \lesssim f'(l), \qquad \text{for } N \geq 6, \\ |l| \lesssim |k| \quad &\Rightarrow \quad \big|f'(k+l) - f'(k) - f'(l)\big| \lesssim |f''(k)| \cdot |l|, \qquad \text{for } N \in \{3,4,5\}. \end{aligned} \tag{3.38}$$

In the region $|x| \leq \sqrt{\lambda}$ we have $W \lesssim W_\lambda$, hence by (3.38)

$$\big|f'(W - W_\lambda) - f'(W_\lambda) - f'(W)\big| \lesssim 1,$$

and $\|1\|_{L^{\frac{N}{2}}(|x|\leq\sqrt{\lambda})} \sim \lambda$.

In the region $|x| \geq \sqrt{\lambda}$ we have $W_\lambda \lesssim W$. If $N \geq 6$, then

$$\big|f'(W - W_\lambda) - f'(W_\lambda) - f'(W)\big| \lesssim f'(W_\lambda).$$

It is easy to check that $\|f'(W_\lambda)\|_{L^{\frac{N}{2}}(|x|\geq\sqrt{\lambda})} \sim \lambda$. If $N \in \{3,4,5\}$, we obtain

$$\big|f'(W - W_\lambda) - f'(W_\lambda) - f'(W)\big| \lesssim |f''(W)| \cdot |W_\lambda|,$$

hence

$$\|f'(W - W_\lambda) - f'(W_\lambda) - f'(W)\|_{L^{\frac{N}{2}}(|x|\geq\sqrt{\lambda})} \lesssim \|f''(W)\|_{L^{\frac{2N}{6-N}}} \cdot \|W_\lambda\|_{L^{\frac{2N}{N-2}}(|x|\geq\sqrt{\lambda})} \sim \lambda^{\frac{N-2}{4}}.$$

This finishes the proof of (3.37).

Step 2. In view of (3.36), it suffices to prove that if $\lambda < \lambda_0$ and $\langle \mathcal{Z}, g\rangle = \langle \mathcal{Z}_{\underline{\lambda}}, g\rangle = 0$, then

$$\frac{1}{2}\langle H_\lambda \boldsymbol{g}, \boldsymbol{g}\rangle + 2\big(\langle \alpha_{\lambda_1}^-, \boldsymbol{g}\rangle^2 + \langle \alpha_{\lambda_1}^+, \boldsymbol{g}\rangle^2 + \langle \alpha_{\lambda_2}^-, \boldsymbol{g}\rangle^2 + \langle \alpha_{\lambda_2}^+, \boldsymbol{g}\rangle^2\big) \gtrsim \|\boldsymbol{g}\|_{\mathcal{E}}^2.$$

Let $a_1^- := \langle \alpha_\lambda^-, \boldsymbol{g}\rangle$, $a_1^+ := \langle \alpha_\lambda^+, \boldsymbol{g}\rangle$, $a_2^- := \langle \alpha^-, \boldsymbol{g}\rangle$, $a_2^+ := \langle \alpha^+, \boldsymbol{g}\rangle$ and decompose

$$\boldsymbol{g} = a_1^- \mathcal{Y}_\lambda^- + a_1^+ \mathcal{Y}_\lambda^+ + a_2^- \mathcal{Y}_\lambda^- + a_2^+ \mathcal{Y}^+ + \boldsymbol{k}.$$

Using the fact that

$$\begin{aligned} |\langle \alpha^\pm, \mathcal{Y}_\lambda^\pm\rangle| + |\langle \alpha_\lambda^\pm, \mathcal{Y}^\pm\rangle| + |\langle \frac{1}{\lambda}\mathcal{Z}_{\underline{\lambda}}, \mathcal{Y}\rangle| + |\langle \mathcal{Z}, \mathcal{Y}_\lambda\rangle| &\lesssim \lambda^{\frac{N-2}{2}}, \\ |a_1^-| + |a_1^+| + |a_2^-| + |a_2^+| &\lesssim \|\boldsymbol{g}\|_{\mathcal{E}}, \\ \langle \alpha^-, \mathcal{Y}^+\rangle = \langle \alpha^+, \mathcal{Y}^-\rangle = \langle \mathcal{Z}, \mathcal{Y}\rangle &= 0 \end{aligned}$$

we obtain

$$\langle \alpha^-, \boldsymbol{k}\rangle^2 + \langle \alpha^+, \boldsymbol{k}\rangle^2 + \langle \alpha_\lambda^-, \boldsymbol{k}\rangle^2 + \langle \alpha_\lambda^+, \boldsymbol{k}\rangle^2 + \langle \mathcal{Z}, k\rangle^2 + \langle \frac{1}{\lambda}\mathcal{Z}_{\underline{\lambda}}, k\rangle^2 \lesssim \lambda^{N-2}\|\boldsymbol{g}\|_{\mathcal{E}}^2. \tag{3.39}$$

Since H_λ is self-adjoint, we can write

$$\begin{aligned} \frac{1}{2}\langle H_\lambda \boldsymbol{g}, \boldsymbol{g}\rangle &= \frac{1}{2}\langle H_\lambda \boldsymbol{k}, \boldsymbol{k}\rangle + \langle H_\lambda(a_2^-\mathcal{Y}^- + a_2^+\mathcal{Y}^+), \boldsymbol{k}\rangle + \langle H_\lambda(a_1^-\mathcal{Y}_\lambda^- + a_1^+\mathcal{Y}_\lambda^+), \boldsymbol{k}\rangle \\ &+ \frac{1}{2}\langle H_\lambda(a_2^-\mathcal{Y}^- + a_2^+\mathcal{Y}^+), a_2^-\mathcal{Y}^- + a_2^+\mathcal{Y}^+\rangle \\ &+ \frac{1}{2}\langle H_\lambda(a_1^-\mathcal{Y}_\lambda^- + a_1^+\mathcal{Y}_\lambda^+), a_1^-\mathcal{Y}_\lambda^- + a_1^+\mathcal{Y}_\lambda^+\rangle \\ &+ \langle H_\lambda(a_2^-\mathcal{Y}^- + a_2^+\mathcal{Y}^+), a_1^-\mathcal{Y}_\lambda^- + a_1^+\mathcal{Y}_\lambda^+\rangle. \end{aligned} \tag{3.40}$$

It is easy to see that $\|f'(W)\mathcal{Y}_\lambda\|_{L^{\frac{2N}{N+2}}} \to 0$ and $\|f'(W_\lambda)\mathcal{Y}\|_{L^{\frac{2N}{N+2}}} \to 0$ as $\lambda \to 0$. This and (2.17), (2.18) imply

$$\|H_\lambda \mathcal{Y}^- + 2\alpha^+\|_{\mathcal{E}^*} + \|H_\lambda \mathcal{Y}^+ + 2\alpha^-\|_{\mathcal{E}^*} + \|H_\lambda \mathcal{Y}_\lambda^- + 2\alpha_\lambda^+\|_{\mathcal{E}^*} + \|H_\lambda \mathcal{Y}_\lambda^+ + 2\alpha_\lambda^-\|_{\mathcal{E}^*} \xrightarrow[\lambda\to 0]{} 0.$$

Plugging this into (3.40) and using (3.39) we obtain

$$\frac{1}{2}\langle H_\lambda \boldsymbol{g}, \boldsymbol{g}\rangle \geq -2a_2^-a_2^+ - 2a_1^-a_1^+ + \frac{1}{2}\langle H_\lambda \boldsymbol{k}, \boldsymbol{k}\rangle - \widetilde{c}\|\boldsymbol{g}\|_{\mathcal{E}}^2, \tag{3.41}$$

where $\widetilde{c} \to 0$ as $\lambda \to 0$.

Applying (2.2) with $r_1 = \lambda^{-\frac{1}{2}}$, rescaling and using (3.39) we get, for λ small enough,

$$(1-2c)\int_{|x|\leq\sqrt{\lambda}} |\nabla k|^2\,\mathrm{d}x + c\int_{|x|\geq\sqrt{\lambda}} |\nabla k|^2\,\mathrm{d}x - \int_{\mathbb{R}^N} f'(W_\lambda)|k|^2\,\mathrm{d}x \geq -\widetilde{c}\|\boldsymbol{g}\|_{\mathcal{E}}^2. \tag{3.42}$$

From (2.3) with $r_2 = \sqrt{\lambda}$ we have

$$(1-2c)\int_{|x|\geq\sqrt{\lambda}} |\nabla k|^2\,\mathrm{d}x + c\int_{|x|\leq\sqrt{\lambda}} |\nabla k|^2\,\mathrm{d}x - \int_{\mathbb{R}^N} f'(W)|k|^2\,\mathrm{d}x \geq -\widetilde{c}\|\boldsymbol{g}\|_{\mathcal{E}}^2. \tag{3.43}$$

Taking the sum of (3.42) and (3.43), and using (3.41) we obtain

$$\frac{1}{2}\langle H_\lambda \boldsymbol{g}, \boldsymbol{g}\rangle \geq -2a_2^-a_2^+ - 2a_1^-a_1^+ + c\|\boldsymbol{k}\|_{\mathcal{E}}^2 - 2\widetilde{c}\|\boldsymbol{g}\|_{\mathcal{E}}^2.$$

The conclusion follows if we take $\widetilde{c}$ small enough. □

Recall that $R > 0$ is the constant used in the definition of the localized bubble $\boldsymbol{V}(\lambda_1, \lambda_2)$, see (3.1).

Lemma 3.6. *There exist constants $\lambda_0, \eta, R_0, c > 0$ such that if $\lambda = \frac{\lambda_1}{\lambda_2} \leq \lambda_0$, $|a_2| \leq \eta$ and $R \geq R_0$, then*

$$E(\boldsymbol{U}(\lambda_1, \lambda_2, a_2)) \geq 2E(\boldsymbol{W}) + c\lambda^{\frac{N-2}{2}}.$$

Proof. Without loss of generality we can assume that $\lambda_2 = 1$, $\lambda_1 = \lambda$ (it suffices to rescale). The conclusion follows from [39, Lemma 2.7] applied for $\boldsymbol{u}^* = -\boldsymbol{U}^{a_2}$ (the proof given there is valid for $N \geq 3$). □

Remark 3.7. In Lemma 3.5 the fact that the bubbles have opposite signs has no importance, but it is crucial in Lemma 3.6.

3.3 Conclusion of the proof

Proof of Theorem 1. Suppose by contradiction that $\boldsymbol{u}(t) : [0, T_+) \to \mathcal{E}$ is a solution of (NLW) such that (1.4) holds. Formula (1.5) and Lemma A.5 imply

$$\begin{aligned} 2E(\boldsymbol{W}) = E(\boldsymbol{U}(\lambda_1, \lambda_2, a_2) + \boldsymbol{g}) &= E(\boldsymbol{U}(\lambda_1, \lambda_2, a_2)) + \langle \mathrm{D}E(\boldsymbol{U}(\lambda_1, \lambda_2, a_2)), \boldsymbol{g}\rangle \\ &\quad + \frac{1}{2}\langle \mathrm{D}^2 E(\boldsymbol{U}(\lambda_1, \lambda_2, a_2))\boldsymbol{g}, \boldsymbol{g}\rangle + o(\|\boldsymbol{g}\|_{\mathcal{E}}^2). \end{aligned} \tag{3.44}$$

Step 1 – Coercivity. We will prove that for all t there holds

$$2a_2(t)b_2(t) + 2\big(a_1^-(t)^2 + a_1^+(t)^2 + b_2(t)^2\big) \gtrsim \mathcal{N}(\boldsymbol{g}(t), \lambda(t))^2 \tag{3.45}$$

(the functions a_1^+, a_1^- and b_2 are defined in Proposition 3.4).

From (2.52) we have $|b_2(t)^2 - \langle \alpha^+_{\lambda_2(t)}, \boldsymbol{g}(t)\rangle^2| \lesssim |a_2| \cdot \|\boldsymbol{g}\|^2$. Since $\langle \alpha^-_{\lambda_2(t)}, \boldsymbol{g}(t)\rangle = 0$, Lemma 3.5 and Lemma 3.6 yield

$$\begin{aligned} E\big(\boldsymbol{U}(\lambda_1(t), \lambda_2(t), a_2(t))\big) - 2E(\boldsymbol{W}) + \frac{1}{2}\big\langle \mathrm{D}^2 E\big(\boldsymbol{U}(\lambda_1(t), \lambda_2(t), a_2(t))\big)\boldsymbol{g}(t), \boldsymbol{g}(t)\big\rangle \\ + 2\big(a_1^-(t)^2 + a_1^+(t)^2 + b_2(t)^2\big) \geq c \cdot \mathcal{N}(\boldsymbol{g}(t), \lambda(t))^2, \end{aligned}$$

for $R \geq R_0$, with a constant $c > 0$ independent of R.

Recall that $2a_2(t)b_2(t) = -\langle \mathrm{D}E(\boldsymbol{U}^{a_2}_{\lambda_2}, \boldsymbol{g}\rangle$. In view of (3.44), in order to prove (3.45) it suffices to verify that

$$|\langle \mathrm{D}E(\boldsymbol{U}(\lambda_1, \lambda_2, a_2)) - \mathrm{D}E(\boldsymbol{U}^{a_2}_{\lambda_2}), \boldsymbol{g}\rangle| \leq \frac{c}{2} \cdot \mathcal{N}(\boldsymbol{g}, \lambda)^2 \tag{3.46}$$

provided that R is large enough. Without loss of generality we can assume that $\lambda_2 = 1$ and $\lambda_1 = \lambda$. First we show that

$$\big|\langle \mathrm{D}E(\boldsymbol{U}(\lambda, 1, a_2)), \boldsymbol{g}\rangle + \langle \mathrm{D}E(\boldsymbol{V}(\lambda, 1)), \boldsymbol{g}\rangle - \langle \mathrm{D}E(\boldsymbol{U}^{a_2}), \boldsymbol{g}\rangle\big| \ll \mathcal{N}(\boldsymbol{g}, \lambda)^2. \tag{3.47}$$

This is equivalent to

$$\int |f(U^{a_2} - V(\lambda, 1)) + f(V(\lambda, 1)) - f(U^{a_2})| \cdot |g| \,\mathrm{d}x \ll \mathcal{N}(\boldsymbol{g}, \lambda)^2.$$

By Hölder and Sobolev inequalities, it suffices to check that

$$\|f(-V(\lambda, 1) + U^{a_2}) + f(V(\lambda, 1)) - f(U^{a_2})\|_{L^{\frac{2N}{N+2}}} \ll \lambda^{\frac{N-2}{4}},$$

which follows from the inequality

$$||f(-V(\lambda,1)+U^{a_2})+f(V(\lambda,1))-f(U^{a_2})|\lesssim f'(W_\lambda)+1.$$

Next, we prove that if R is large enough, then

$$\|\mathrm{D}E(\boldsymbol{V}(\lambda,1))\|_{\mathcal{E}^*}\leq\frac{c}{4}\cdot\lambda^{\frac{N-2}{4}}. \tag{3.48}$$

From (3.3), if R is large then

$$\|\Delta(W_\lambda-V(\lambda,1))\|_{\dot H^{-1}}\lesssim\frac{c}{8}\cdot\lambda^{\frac{N-2}{4}}. \tag{3.49}$$

We will prove that

$$\|f(W_\lambda)-f(V(\lambda,1))\|_{L^{\frac{2N}{N+2}}}\ll\lambda^{\frac{N-2}{4}}. \tag{3.50}$$

In the region $|x|\geq R\sqrt{\lambda}$ we have $V(\lambda,1)=0$ and

$$\|f(W_\lambda)\|_{L^{\frac{2N}{N+2}}(|x|\geq R\sqrt{\lambda})}=\|f(W)\|_{L^{\frac{2N}{N+2}}(|x|\geq R/\sqrt{\lambda})}\sim\lambda^{\frac{N+2}{4}}\ll\lambda^{\frac{N-2}{2}}.$$

In the region $|x|\leq R\sqrt{\lambda}$ we use the pointwise bound $|f(W_\lambda)-f(V(\lambda,1))|\lesssim f'(W_\lambda)\cdot|W_\lambda-V(\lambda,1)|$, the fact that $W_\lambda-V(\lambda,1)$ is bounded in L^∞ and the bound

$$\|f'(W_\lambda)\|_{L^{\frac{2N}{N+2}}(|x|\leq R\sqrt{\lambda})}\ll\lambda^{\frac{N-2}{4}}.$$

Now (3.48) follows from (3.49), (3.50) and $\Delta W_\lambda+f(W_\lambda)=0$.

Estimate (3.46) follows from (3.47) and (3.48).

Step 2 – Differential inequalities. Observe that

$$\int_0^{T_+}\frac{1}{\lambda_1(t)}\,\mathrm{d}t=\int_0^{T_+}\frac{1}{\lambda_2(t)}\,\mathrm{d}t=+\infty. \tag{3.51}$$

The proof is the same as the proof of (2.64).

For $m\in\mathbb{N}$, $m\geq m_0$, let $t=t_m$ be the last time such that $\mathcal{N}(\boldsymbol{g}(t),\lambda(t))=2^{-m}$. By continuity, t_m is well defined if m_0 is large enough.

By Proposition 3.4, there exists a constant C_1 such that

$$|a_1^+(t)|\geq C_1\cdot\mathcal{N}(\boldsymbol{g}(t),\lambda(t))\quad\Rightarrow\quad\frac{\mathrm{d}}{\mathrm{d}t}|a_1^+(t)|\geq\frac{\nu}{2\lambda_1(t)}|a_1^+(t)|,\qquad\forall t\in[0,T_+). \tag{3.52}$$

Suppose that $|a_1^+(t_m)|\geq 2C_1\cdot\mathcal{N}(\boldsymbol{g}(t_m),\lambda(t_m))$. Since, by the definition of t_m, $\mathcal{N}(\boldsymbol{g}(t),\lambda(t))\leq\mathcal{N}(\boldsymbol{g}(t_m),\lambda(t_m))$ for $t\geq t_m$, a simple continuity argument yields $|a_1^+(t_m)|\geq 2C_1\cdot\mathcal{N}(\boldsymbol{g}(t),\lambda(t))$ for all $t\geq t_m$. By (3.52) and (3.51), this implies $|a_1^+(t)|\to+\infty$ as $t\to T_+$, which is absurd. The same reasoning applies to $b(t)$, hence we get

$$|a_1^+(t_m)|\lesssim\mathcal{N}(\boldsymbol{g}(t_m),\lambda(t_m))^2,\qquad|b(t_m)|\lesssim\mathcal{N}(\boldsymbol{g}(t_m),\lambda(t_m))^2.$$

Thus (3.45) forces

$$|a_1^-(t_m)|\gtrsim\mathcal{N}(\boldsymbol{g}(t_m),\lambda(t_m))\gg\mathcal{N}(\boldsymbol{g}(t_m),\lambda(t_m))^2. \tag{3.53}$$

Consider the evolution on the time interval $[t_{m-1}, t_m]$. By definition of t_{m-1} and t_m for $t \in [t_{m-1}, t_m]$ there holds $\mathcal{N}(\boldsymbol{g}(t), \lambda(t)) \leq 2\cdot\mathcal{N}(\boldsymbol{g}(t_m), \lambda(t_m))$, hence (3.53) and Proposition 3.4 allow to conclude that

$$\frac{\mathrm{d}}{\mathrm{d}t}|a_1^-(t)| \leq -\frac{\nu}{2\lambda_1(t)}|a_1^-(t)|, \qquad \forall t \in [t_{m-1}, t_m].$$

Since this holds for all m sufficiently large, we deduce that there exists $t_0 < T_+$ such that

$$|a_1^-(t)| \leq |a_1^-(t_0)| \cdot \exp\Big(-\int_{t_0}^t \frac{\nu\,\mathrm{d}\tau}{2\lambda_1(\tau)}\Big), \qquad \forall t \geq t_0.$$

Let $t \in [t_{m-1}, t_m]$. At time t_m all the terms of (3.45) except for the term $2a_1^-(t)^2$ are absorbed by the right hand side, hence $\mathcal{N}(\boldsymbol{g}(t_m), \lambda(t_m)) \lesssim |a_1^-(t_m)|$. Using the definition of t_{m-1} we obtain

$$\begin{aligned}\mathcal{N}(\boldsymbol{g}(t), \lambda(t)) \leq 2\mathcal{N}(\boldsymbol{g}(t_m), \lambda(t_m)) \lesssim |a_1^-(t_m)| \lesssim |a_1^-(t_0)| \cdot \exp\Big(-\int_{t_0}^{t_m} \frac{\nu\,\mathrm{d}\tau}{2\lambda_1(\tau)}\Big) \\ \lesssim |a_1^-(t_0)| \cdot \exp\Big(-\int_{t_0}^{t} \frac{\nu\,\mathrm{d}\tau}{2\lambda_1(\tau)}\Big).\end{aligned}$$

By (3.14), this implies

$$|\lambda_1'(t)| + |\lambda_2'(t)| \lesssim \exp\Big(-\int_{t_0}^t \frac{\nu\,\mathrm{d}\tau}{2\lambda_1(\tau)}\Big), \qquad \forall t \geq t_0.$$

Dividing both sides by $\lambda_1(t)$ and integrating we obtain that $\log\lambda_1(t)$ converges as $t \to T_+$. Dividing both sides by $\lambda_2(t)$, using the fact that $\lambda_2(t) \geq \lambda_1(t)$ for $t \geq t_0$ and integrating we obtain that $\log\lambda_2(t)$ converges as $t \to T_+$. Hence $\log\lambda(t)$ converges, which is impossible. □

Remark 3.8. An analogous proof using the linear stability and instability components $\alpha_{\lambda_2}^+$ and $\alpha_{\lambda_2}^-$ instead of the refined modulation and instability component $\beta_{\lambda_2}^{a_2}$ would yield $\lambda_2(0) \to \lambda_0 \in (0, +\infty)$ (hence $T_+ = +\infty$) and $|\log\lambda_1(t)| \gtrsim t$ as $t \to +\infty$, but would not (at least directly) lead to a contradiction.

A Elementary lemmas

Lemma A.1. *Let $\psi : \mathbb{R} \to \mathbb{R}$ be an analytic function such that $\psi(0) = 0$ and $\psi'(0) \neq 0$. Then there exists a local analytic diffeomorphism $y = \varphi(x)$ near $x = 0$ such that $\varphi(0) = 0$, $\varphi'(0) = 1$ and*

$$\varphi'(x) \cdot \psi(x) = \varphi(x) \cdot \psi'(0). \tag{A.1}$$

Remark A.2. Equation (A.1) expresses the fact that the change of variable $y = \varphi(x)$ transforms the differential equation $\dot{x} = \psi(x)$ into $\dot{y} = \psi'(0)y$.

Proof. Without loss of generality we can assume that $\psi'(0) = 1$. We set:

$$\varphi(x) := \psi(x) \cdot \exp\Big(\int_0^x \frac{1 - \psi'(z)}{\psi(z)}\,\mathrm{d}z\Big)$$

and it suffices to verify that φ has the required properties. □

Recall that we denote $f(u) := |u|^{\frac{4}{N-2}}u$ and $R(v) := f(W + v) - f(W) - f'(W)v$. Notice that f' is not Lipshitz for $N > 6$.

Lemma A.3. *The mapping $\mathcal{R}$ is analytic from $B_{Y^k}(0,\eta)$ to itself if $k \geq k_0$ and η is small. Its derivative is given by*

$$\mathscr{L}(Y^k) \ni \mathrm{D}_v\mathcal{R} = \big(h \mapsto (f'(W+v) - f'(W))h\big).$$

The same conclusion holds if we replace Y^k by $BC_{\widetilde{\nu}}$ for $\widetilde{\nu} \geq 0$.

Proof. We have an isomorphism

$$\Phi : Y^k \to H^k, \qquad \Phi(v) := (1+|x|^k)v,$$

so it suffices to show that $\Phi \circ \mathcal{R} \circ \Phi^{-1}$ is analytic from $B_{H^k}(0,\eta)$ to itself. Let $w \in B_{H^k}(0,\eta)$.

Let $f(1+z) = |1+z|^{\frac{4}{N}}(1+z) = \sum_{n=0}^{+\infty} a_n z^n$. The series converges for $|z| < 1$. We have a series expansion:

$$\mathcal{R}(\Phi^{-1}w) = \sum_{n=2}^{+\infty} a_n W^{\frac{N+2}{N-2}-n}\frac{w^n}{(1+|x|^k)^n} = \frac{1}{1+|x|^k}\frac{W^{\frac{6-N}{N-2}}}{1+|x|^k}\sum_{n=2}^{+\infty} a_n\Big(\frac{1}{W\cdot(1+|x|^k)}\Big)^{n-2} w^n.$$

We see that $\frac{W^{\frac{6-N}{N-2}}}{1+|x|^k} \in H^k$ if k is large enough and that the last series converges strongly in H^k if η is small.

In the case of the space $BC_{\widetilde{\nu}}$ the proof is the same. □

Lemma A.4. *There exists $k = k(N) \in \mathbb{N}$ and $\eta = \eta(N) > 0$ such that if $\psi \in Y^k$ and $|a| \leq \eta$, then for all $g \in \dot H^1$ such that $\|g\|_{\dot H^1} \leq \eta$ there holds*

$$|\langle \psi, f(U^a+g) - f(U^a) - f'(U^a)g\rangle| \lesssim \|g\|_{\dot H^1}^2,$$
$$|\langle \psi, \big(f'(U^a) - f'(W)\big)g\rangle| \lesssim |a|\cdot\|g\|_{\dot H^1}, \tag{A.2}$$

with a constant depending on ψ.

Proof. For $N \in \{3,4,5\}$ this follows directly from the Sobolev and Hölder inequalities (even for $\psi \in \dot H^1$).

For $N \geq 6$ we use the pointwise bound

$$|f(U^a+g) - f(U^a) - f'(U^a)g| \lesssim |f''(U^a)|\cdot|g|^2.$$

Here, f'' is a negative power. Since U^a has slow decay, $\psi\cdot|f''(U^a)| \in L^{\frac{N}{2}}$ if $\psi \in Y^k$ and k is large enough. The conclusion follows from the Hölder inequlity.

The proof of (A.2) is similar. □

Lemma A.5. *Let $\gamma := \min\big(3, \frac{2N}{N-2}\big)$. For any $M > 0$ there exists $C > 0$ and $\eta > 0$ such that if $\|\boldsymbol{v}\|_{\mathcal{E}} \leq M$ and $\|\boldsymbol{g}\|_{\mathcal{E}} \leq \eta$, then*

$$\big|E(\boldsymbol{v}+\boldsymbol{g}) - E(\boldsymbol{v}) - \langle \mathrm{D}E(\boldsymbol{v}),\boldsymbol{g}\rangle - \frac12\langle \mathrm{D}^2E(\boldsymbol{v})\boldsymbol{g},\boldsymbol{g}\rangle\big| \leq C\|\boldsymbol{g}\|_{\mathcal{E}}^\gamma.$$

Proof. In dimension $N \in \{3,4,5\}$ this follows from the pointwise inequality

$$\big|F(k+l) - F(k) - f(k)l - \frac12 f'(k)l^2\big| \lesssim |f''(k)|\cdot|l^3| + |F(l)|, \qquad k,l \in \mathbb{R}, \tag{A.3}$$

whereas for $N \geq 6$ from

$$\big|F(k+l) - F(k) - f(k)l - \frac12 f'(k)l^2\big| \lesssim |F(l)|, \qquad k,l \in \mathbb{R}. \tag{A.4}$$

In order to prove bounds (A.3) and (A.4), notice that they are homogeneous and invariant by changing signs of both k and l, hence it can be assumed that $k = 1$ (for $k = 0$ the inequalities are obvious). Now for $|l| \leq \frac12$ the conclusion follows from the asymptotic expansion of $F(1+l)$ and for $|l| \geq \frac12$ the bounds are evident. □

Bibliography

[1] T. Aubin. Équations différentielles non linéaires et problème de Yamabe concernant la courbure scalaire. *J. Math. Pures Appl.*, 55(9):269–296, 1976.

[2] H. Bahouri, J.-Y. Chemin, and Danchin R. *Fourier Analysis and Nonlinear Partial Differential Equations*, volume 343 of *Grundlehren der mathematischen Wissenschaften.* Springer, 2011.

[3] H. Bahouri and P. Gérard. High frequency approximation of solutions to critical nonlinear wave equations. *Amer. J. Math.*, 121(1):131–175, 1999.

[4] P. Bizoń, T. Chmaj, and Z. Tabor. Formation of singularities for equivariant 2 + 1 dimensional wave maps into the 2-sphere. *Nonlinearity*, 14(5):1041–1053, 2001.

[5] J. Bourgain and W. Wang. Construction of blowup solutions for the nonlinear Schrödinger equation with critical nonlinearity. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 25:197–215, 1997.

[6] P. Brenner. On scattering and everywhere defined scattering operators for nonlinear Klein-Gordon equations. *J. Differential Equations*, 56(3):310–344, 1985.

[7] F. E. Browder. On non-linear wave equations. *Math. Zeitschr.*, 80:249–264, 1962.

[8] A. Bulut, M. Czubak, D. Li, N. Pavlović, and X. Zhang. Stability and unconditional uniqueness of solutions for energy critical wave equations in high dimensions. *Comm. Part. Diff. Eq.*, 38(4):575–607, 2013.

[9] T. Cazenave. *Semilinear Schrödinger Equations*, volume 10 of *Courant Lecture Notes in Mathematics.* AMS, 2003.

[10] T. Cazenave, J. Shatah, and A. Tahvildar-Zadeh. Harmonic maps of the hyperbolic space and development of singularities in wave maps and Yang-Mills fields. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 68(3):315–349, 1998.

[11] S.-N. Chow and J. K. Hale. *Methods of Bifurcation Theory*, volume 251 of *Grundlehren der mathematischen Wissenschaften.* Springer, 1982.

[12] F. M. Christ and M. I. Weinstein. Dispersion of small amplitude solutions of the generalized Korteweg-de Vries equation. *J. Funct. Anal.*, 100:87–109, 1991.

[13] C. Collot. Type II blow up for the energy supercritical wave equation. *Preprint*, arXiv:1407.4525, 2014.

[14] R. Côte. On the soliton resolution for equivariant wave maps to the sphere. *Comm. Pure Appl. Math.*, 68(11):1946–2004, 2015.

[15] R. Côte, C. Kenig, A. Lawrie, and W. Schlag. Profiles for the radial focusing 4d energy-critical wave equation. *Preprint*, arXiv:1402.2307, 2014.

[16] R. Côte, C. E. Kenig, A. Lawrie, and W. Schlag. Characterization of large energy solutions of the equivariant wave map problem: I. *Amer. J. Math.*, 137(1):139–207, 2015.

[17] R. Côte, C. E. Kenig, and F. Merle. Scattering below critical energy for the radial 4D Yang-Mills equation and for the 2D corotational wave map system. *Comm. Math. Phys.*, 284(1):203–225, 2008.

[18] R. Côte, Y. Martel, and F. Merle. Construction of multi-soliton solutions for the L^2-supercritical gKdV and NLS equations. *Rev. Mat. Iberoam.*, 27(1):273–302, 2011.

[19] R. Donninger, M. Huang, J. Krieger, and W. Schlag. Exotic blowup solutions for the u^5 focusing wave equation in $\mathbb{R}^3$. *Michigan Math. J.*, 63(3):451–501, 2014.

[20] R. Donninger and J. Krieger. Nonscattering solutions and blowup at infinity for the critical wave equation. *Math. Ann.*, 357(1):89–163, 2013.

[21] R. Donninger and B. Schörkhuber. Stable blow up dynamics for energy supercritical wave equations. *Trans. Amer. Math. Soc.*, 366(4):2167–2189, 2014.

[22] R. Donninger and B. Schörkhuber. Stable blowup for wave equations in odd space dimensions. *Preprint*, arXiv:1504.00808, 2015.

[23] T. Duyckaerts, C. E. Kenig, and F. Merle. Universality of blow-up profile for small radial type II blow-up solutions of the energy-critical wave equation. *J. Eur. Math. Soc.*, 13(3):533–599, 2011.

[24] T. Duyckaerts, C. E. Kenig, and F. Merle. Profiles of bounded radial solutions of the focusing, energy-critical wave equation. *Geom. Funct. Anal.*, 22(3):639–698, 2012.

[25] T. Duyckaerts, C. E. Kenig, and F. Merle. Universality of the blow-up profile for small type II blow-up solutions of the energy-critical wave equation: the nonradial case. *J. Eur. Math. Soc.*, 14(5):1389–1454, 2012.

[26] T. Duyckaerts, C. E. Kenig, and F. Merle. Classification of the radial solutions of the focusing, energy-critical wave equation. *Camb. J. Math.*, 1(1):75–144, 2013.

[27] T. Duyckaerts and F. Merle. Dynamics of threshold solutions for energy-critical wave equation. *Int. Math. Res. Pap. IMRP*, 2008.

[28] W. Eckhaus and P. C. Schuur. The emergence of solitons of the Korteweg-de Vries equation from arbitrary initial conditions. *Math. Methods Appl. Sci.*, 5:97–116, 1983.

[29] L. C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. AMS, 1998.

[30] K. O. Friedrichs. Symmetric hyperbolic linear differential equations. *Comm. Pure Appl. Math.*, 7:345–392, 1954.

[31] Y. Giga and R. V. Kohn. Nondegeneracy of blowup for semilinear heat equations. *Comm. Pure Appl. Math.*, 42(6):845–884, 1989.

[32] J. Ginibre, A. Soffer, and G. Velo. The global Cauchy problem for the critical nonlinear wave equation. *J. Funct. Anal.*, 110:96–130, 1992.

[33] J. Ginibre and G. Velo. Generalized Strichartz inequalities for the wave equation. *J. Funct. Anal.*, 133:50–68, 1995.

[34] M. Grillakis. Regularity and asymptotic behavior of the wave equation with a critical nonlinearity. *Ann. of Math.*, 132:485–509, 1990.

[35] S. Gustafson, K. Kang, and T.-P. Tsai. Schrödinger flow near harmonic maps. *Comm. Pure Appl. Math.*, 60(4):463–499, 2007.

[36] M. Hillairet and P. Raphaël. Smooth type II blow up solutions to the four dimensional energy critical wave equation. *Anal. PDE*, 5(4):777–829, 2012.

[37] R. van der Hout. On the nonexistence of finite time bubble trees in symmetric harmonic map heat flows from the disk to the 2-sphere. *J. Differential Equations*, 192(1):188–201, 2003.

[38] S. Ibrahim, N. Masmoudi, and K. Nakanishi. Scattering threshold for the focusing nonlinear Klein-Gordon equation. *Anal. PDE*, 4(3):405–460, 2011.

[39] J. Jendrej. Bounds on the speed of type II blow-up for the energy critical wave equation in the radial case. *Int. Math. Res. Not.*, 2015.

[40] J. Jendrej. Construction of type II blow-up solutions for the energy-critical wave equation in dimension 5. *Preprint*, arXiv:1503.05024, 2015.

[41] J. Jendrej. Nonexistence of radial two-bubbles with opposite signs for the energy-critical wave equation. *Preprint*, arXiv:1510.03965, 2015.

[42] H. Jia and C. E. Kenig. Asymptotic decomposition for semilinear wave and equivariant wave map equations. *Preprint*, arXiv:1503.06715, 2015.

[43] K. Jörgens. Das Anfangswertproblem im Großen für eine Klasse nichtlinearer Wellengleichungen. *Math. Zeitschr.*, 77:295–308, 1961.

[44] L. V. Kapitanski. Global and unique weak solutions of nonlinear wave equations. *Math. Res. Lett.*, 1:211–223, 1994.

[45] J. B. Keller. On solutions of nonlinear wave equations. *Comm. Math. Pure Appl.*, 10:523–530, 1957.

[46] C. E. Kenig and F. Merle. Global well-posedness, scattering and blow-up for the energy-critical, focusing, non-linear schrödinger equation in the radial case. *Invent. Math.*, 166(3):645–675, 2006.

[47] C. E. Kenig and F. Merle. Global well-posedness, scattering and blow-up for the energy-critical focusing non-linear wave equation. *Acta Math.*, 201(2):147–212, 2008.

[48] C. E. Kenig and F. Merle. Nondispersive radial solutions to energy supercritical non-linear wave equations, with applications. *Amer. J. Math.*, 133(4):1029–1065, 2011.

[49] J. Krieger, K. Nakanishi, and W. Schlag. Center-stable manifold of the ground state in the energy space for the critical wave equation. *Math. Ann.*, 361(1–2):1–50, 2015.

[50] J. Krieger and W. Schlag. Full range of blow up exponents for the quintic wave equation in three dimensions. *J. Math Pures Appl.*, 101(6):873–900, 2014.

[51] J. Krieger and W. Schlag. Large global solutions for energy supercritical nonlinear wave equations on $\mathbb{R}^{3+1}$. *J. Anal. Math.*, accepted.

[52] J. Krieger, W. Schlag, and D. Tataru. Renormalization and blow up for charge one equivariant critical wave maps. *Invent. Math.*, 171(3):543–615, 2008.

[53] J. Krieger, W. Schlag, and D. Tataru. Slow blow-up solutions for the $H^1(\mathbb{R}^3)$ critical focusing semilinear wave equation. *Duke Math. J.*, 147(1):1–53, 2009.

[54] A. Lawrie and S.-J Oh. A refined threshold theorem for (1+2)-dimensional wave maps into surfaces. *Comm. Math. Phys.*, 342(3):989–999, 2016.

[55] Y. Y. Li. Prescribing scalar curvature on S^n and related problems, I. *J. Differ. Equ.*, 120(2):319–410, 1995.

[56] Y. Martel. Asymptotic N-soliton-like solutions of the subcritical and critical generalized Korteweg-de Vries equations. *Amer. J. Math.*, 127(5):1103–1140, 2005.

[57] Y. Martel and F. Merle. Instability of solitons for the critical generalized Korteweg-de Vries equation. *Geom. Funct. Anal.*, 11:74–123, 2001.

[58] Y. Martel and F. Merle. Construction of multi-solitons for the energy-critical wave equation in dimension 5. *Preprint*, arXiv:1504.01595, 2015.

[59] Y. Martel, F. Merle, and P. Raphaël. Blow up for the critical gKdV equation I: dynamics near the soliton. *Acta Math.*, 212(1):59–140, 2014.

[60] Y. Martel, F. Merle, and P. Raphaël. Blow up for the critical gKdV equation III: exotic regimes. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, XIV:575–631, 2015.

[61] Y. Martel, F. Merle, P. Raphaël, and J. Szeftel. Near soliton dynamics and singularity formation for L^2 critical problems. *Russ. Math. Surv.*, 69(2):261–290, 2014.

[62] Y. Martel and P. Raphaël. Strongly interacting blow up bubbles for the mass critical NLS. *Preprint*, arXiv:1512.00900, 2015.

[63] H. Matano and F. Merle. Classification of type I and type II behaviors for a supercritical nonlinear heat equation. *J. Funct. Anal.*, 256(4):992–1064, 2009.

[64] F. Merle. Construction of solutions with exactly k blow-up points for the Schrödinger equation with critical nonlinearity. *Commun. Math. Phys.*, 129(2):223–240, 1990.

[65] F. Merle and P. Raphaël. Profiles and quantization of the blow up mass for critical nonlinear Schrödinger equation. *Commun. Math. Phys.*, 253(3):675–704, 2005.

[66] F. Merle, P. Raphaël, and I. Rodnianski. Blowup dynamics for smooth data equivariant solutions to the critical Schrödinger map problem. *Invent. Math.*, 193(2):249–365, 2013.

[67] F. Merle and H. Zaag. Determination of the blow-up rate for the semilinear wave equation. *Amer. J. Math.*, 125(5):1147–1164, 2003.

[68] F. Merle and H. Zaag. Determination of the blow-up rate for a critical semilinear wave equation. *Math. Ann.*, 331(2):395–416, 2005.

[69] N. Mizoguchi. Rate of Type II blowup for a semilinear heat equation. *Math. Ann.*, 339(4):839–877, 2007.

[70] C. S. Morawetz and W. A. Strauss. Decay and scattering of solutions of a nonlinear relativistic wave equation. *Comm. Pure Appl. Math.*, 25(1):1–31, 1972.

[71] K. Nakanishi and W. Schlag. Global dynamics above the ground state energy for the focusing nonlinear Klein-Gordon equation. *J. Differential Equations*, 250(5):2299–2333, 2011.

[72] K. Nakanishi and W. Schlag. Global dynamics above the ground state for the nonlinear Klein-Gordon equation without a radial assumption. *Arch. Ration. Mech. Anal.*, 203(3):809–851, 2011.

[73] M. del Pino, M. Musso, F. Pacard, and A. Pistoia. Large energy entire solutions for the Yamabe equation. *J. Differ. Equ.*, 251(9):2568–2597, 2011.

[74] C. Ortoleva and G. Perelman. Nondispersive vanishing and blow up at infinity for the energy critical nonlinear Schrödinger equation in $\mathbb{R}^3$. *Algebra i Analiz*, 25(2):162–192, 2013.

[75] L. E. Payne and D. H. Sattinger. Saddle points and instability of nonlinear hyperbolic equations. *Israel J. Math.*, 22(3):273–303, 1975.

[76] G. Perelman. Blow up dynamics for equivariant critical Schrödinger maps. *Commun. Math. Phys.*, 330(1):69–105, 2014.

[77] A. Pistoia and T. Weth. Sign changing bubble tower solutions in a slightly subcritical semilinear Dirichlet problem. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 24(2):325–340, 2007.

[78] P. Raphaël and J. Szeftel. Existence and uniqueness of minimal mass blow up solutions to an inhomogeneous L^2-critical NLS. *J. Amer. Math. Soc.*, 24(2):471–546, 2011.

[79] I. Rodnianski and P. Raphaël. Stable blow up dynamics for critical corotational wave maps and the equivariant Yang Mills problem. *Publ. Math. Inst. Hautes Études Sci.*, 115:1–122, 2012.

[80] I. Rodnianski and J. Sterbenz. On the formation of singularities in the critical $O(3)$ σ-model. *Ann. of Math.*, 172(1):187–242, 2010.

[81] C. Rodriguez. Profiles for the radial focusing energy-critical wave equation in odd dimensions. *Preprint*, arXiv:1412.1388, 2014.

[82] R. Schweyer. Type II blow-up for the four dimensional energy critical semi linear heat equation. *J. Funct. Anal.*, 263(12):3922–3983, 2012.

[83] I. E. Segal. Non-linear semi-groups. *Ann. of Math.*, 78(2):339–364, 1963.

[84] J. Shatah and M. Struwe. Well-posedness in the energy space for semilinear wave equations with critical growth. *Internat. Math. Res. Notices*, 7:303–309, 1994.

[85] J. Shatah and M. Struwe. *Geometric Wave Equations*, volume 2 of *Courant Lecture Notes in Mathematics.* AMS, 2000.

[86] J. Shatah and A. Tahvildar-Zadeh. On the Cauchy problem for equivariant wave maps. *Comm. Pure Appl. Math.*, 47(5):719–754, 1994.

[87] W. A. Strauss. Existence of solitary waves in higher dimensions. *Comm. Math. Phys.*, 55:149–162, 1977.

[88] M. Struwe. Globally regular solutions to the u^5 klein-gordon equation. *Ann. Scuola Norm. Pisa Cl. Sci.*, 15:495–513, 1988.

[89] M. Struwe. Equivariant wave maps in two space dimensions. *Comm. Pure Appl. Math.*, 56(7):815–823, 2003.

[90] G. Talenti. Best constant in Sobolev inequality. *Ann. Mat. Pura Appl.*, 110(4):353–372, 1976.

[91] T. Tao. *Nonlinear Dispersive Equations: Local and Global Analysis*, volume 106 of *CBMS Reg. Conf. Series in Math.* AMS, 2006.

[92] G. Teschl. *Ordinary Differential Equations and Dynamical Systems*, volume 140 of *Graduate Studies in Mathematics.* AMS, 2012.

[93] P. Topping. An example of a nontrivial bubble tree in the harmonic map heat flow. In *Harmonic Morphisms, Harmonic Maps and Related Topics.* Chapman and Hall/CRC, 1999.

[94] P. Topping. Repulsion and quantization in almost-harmonic maps, and asymptotics of the harmonic map flow. *Ann. of Math.*, 159(2):465–534, 2004.

Université Paris-Saclay | École doctorale de mathématiques Hadamard (EDMH)

Titre : Sur la dynamique d'équations des ondes avec non-linéarité énergie critique focalisante

Mots clés : équation des ondes, non linéarité énergie-critique, explosion, multi-soliton

Résumé : Cette thèse est consacrée à l'étude du comportement global des solutions de l'équation des ondes énergie-critique focalisante. On s'intéresse tout spécialement à la description de la dynamique du système dans l'espace d'énergie. Nous développons une variante de la méthode d'énergie qui permet de construire des solutions explosives de type II, instables. Ensuite, par une démarche similaire, nous donnons le premier exemple d'une solution radiale de l'équation des ondes énergie-critique qui converge dans l'espace d'énergie vers une superposition de deux états stationnaires (bulles). En appliquant notre méthode au cas de l'équation des ondes des applications harmoniques (wave map), nous obtenons des solutions de type bulle-antibulle, en toute classe d'équivariance $k > 2$. Pour l'équation des ondes énergie-critique radiale, nous étudions également le lien entre la vitesse de l'explosion de type II et la limite faible de la solution au moment de l'explosion. Finalement, nous montrons qu'il est impossible qu'une solution radiale converge vers une superposition de deux bulles ayant les signes opposés.

Title : On the dynamics of energy-critical focusing wave equations

Keywords : wave equation, energy-critical nonlinearity, blow-up, multisoliton

Abstract : In this thesis we study the global behavior of solutions of the energy-critical focusing nonlinear wave equation, with a special emphasis on the description of the dynamics in the energy space. We develop a new approach, based on the energy method, to constructing unstable type II blow-up solutions. Next, we give the first example of a radial two-bubble solution of the energy-critical wave equation. By implementing this construction in the case of the equivariant wave map equation, we obtain bubble-antibubble solutions in equivariance classes $k > 2$. We also study the relationship between the speed of a type II blow-up and the weak limit of the solution at the blow-up time. Finally, we prove that there are no pure radial two-bubbles with opposite signs for the energy-critical wave equation.

École Polytechnique
Université Paris-Saclay

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France