# The monoatomic FPU system as a limit of a diatomic FPU system 

Dmitry E. Pelinovsky ${ }^{\text {a,b,* }}$, Guido Schneider ${ }^{\text {c }}$<br>${ }^{\text {a }}$ Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada, L8S 4K1<br>${ }^{\mathrm{b}}$ Department of Applied Mathematics, Nizhny Novgorod State Technical University, Nizhny Novgorod 603950, Russia<br>${ }^{\text {c }}$ Institut für Analysis, Dynamik und Modellierung, Universität Stuttgart, Pfaffenwaldring 57, D-70569 Stuttgart, Germany

## A R T I C L E I N F O

Article history:
Received 28 February 2020
Received in revised form 7 April 2020
Accepted 7 April 2020
Available online 18 April 2020

## Keywords:

Fermi-Pasta-Ulam system
The small mass ratio limit
Justification of approximations
Energy estimates


#### Abstract

We consider a diatomic infinite Fermi-Pasta-Ulam (FPU) system with light and heavy particles. For a small mass ratio, we prove error estimates for the approximation of the dynamics of this system by the dynamics of the monoatomic FPU system. The light particles are squeezed by the heavy particles at the average value of their displacements. The error estimates are derived by means of the energy method and hold for sufficiently long times, for which the dynamics of the monoatomic FPU system is observed. The approximation result is restricted to sufficiently small displacements of the heavy particles relatively to each other.


© 2020 Elsevier Ltd. All rights reserved.

## 1. Introduction

We consider a diatomic infinite Fermi-Pasta-Ulam (FPU) system depicted schematically in Fig. 1. Displacements of heavy particles are denoted by $Q_{j}$ with $j \in 2 \mathbb{Z}$, whereas displacements of light particles are denoted by $q_{j}$, with $j \in 2 \mathbb{Z}+1$. For convenience, we normalize the mass of the heavy particles to unity and denote the mass ratio between masses of light and heavy particles by the parameter $\varepsilon^{2}$. The total energy of the diatomic system is

$$
\begin{equation*}
H=\sum_{j \in 2 \mathbb{Z}} \frac{1}{2} \dot{Q}_{j}^{2}+\frac{1}{2} \varepsilon^{2} \dot{q}_{j+1}^{2}+W\left(q_{j+1}-Q_{j}\right)+W\left(Q_{j}-q_{j-1}\right), \tag{1}
\end{equation*}
$$

where the dot denotes the derivative in time $t$ and $W: \mathbb{R} \mapsto \mathbb{R}$ is a smooth potential for the pairwise interaction force between the adjacent light and heavy particles. Equations of motion are generated from

[^0]

Fig. 1. A diatomic FPU system with heavy and light particles.
the total energy (1) by using the standard symplectic structure for the dynamics of particles. They are written in the form:

$$
\begin{align*}
\ddot{Q}_{j} & =W^{\prime}\left(q_{j+1}-Q_{j}\right)-W^{\prime}\left(Q_{j}-q_{j-1}\right)  \tag{2}\\
\varepsilon^{2} \ddot{q}_{j+1} & =W^{\prime}\left(Q_{j+2}-q_{j+1}\right)-W^{\prime}\left(q_{j+1}-Q_{j}\right) \tag{3}
\end{align*}
$$

where $j \in 2 \mathbb{Z}$.
The dynamics of diatomic lattices, e.g. propagation of traveling solitary waves, has always been important in physical applications and has been studied in numerous works, e.g. [1-3]. More recently, such diatomic systems were considered in the context of granular chains [4-6]. In particular, the authors of [4] proposed to consider the following reduction of the diatomic system in the limit of vanishing mass ratio $\varepsilon \rightarrow 0$ :

$$
0=W^{\prime}\left(Q_{j+2}-q_{j+1}\right)-W^{\prime}\left(q_{j+1}-Q_{j}\right) \quad \Rightarrow \quad Q_{j+2}-q_{j+1}=q_{j+1}-Q_{j}
$$

which yields

$$
\begin{equation*}
q_{j+1}=\frac{Q_{j+2}+Q_{j}}{2} \tag{4}
\end{equation*}
$$

If $q_{j+1}$ is expressed by (4), the dynamics of the heavy particles is governed by the monoatomic FPU system:

$$
\begin{equation*}
\ddot{Q}_{j}=W^{\prime}\left(\frac{Q_{j+2}-Q_{j}}{2}\right)-W^{\prime}\left(\frac{Q_{j}-Q_{j-2}}{2}\right) \tag{5}
\end{equation*}
$$

where $j \in 2 \mathbb{Z}$. It follows from (4) and (5) that the light particles are squeezed by the heavy particles and move according to the average value of the displacements of their heavy neighbors, whereas the heavy particles move according to their pairwise interactions.

Numerical results on existence and non-existence of traveling solitary waves in the diatomic system (2)(3) which are close to the traveling solitary waves of the monoatomic system (5) were reported in [4]. These numerical results inspired a number of analytical works where the authors developed the existence theory for traveling solitary waves with oscillatory tails $[7,8]$, beyond-all-order theory $[9,10]$, and the linearized analysis of perturbations [11]. It is the purpose of this paper to give rigorous error estimates for this approximation in the context of the initial-value problem.

Note that the small mass ratio limit for diatomic FPU system has been considered before in the context of the existence of breathers [12-15] and traveling periodic waves [16-19]. However, these works rely on the ideas of the so-called anti-continuum limit, for which the heavy particles do not move after rescaling of the time variable, whereas the light particles perform uncoupled oscillations in between the heavy particles. The limit (4) and (5) is clearly different from the anti-continuum limit.

Other relevant results on traveling solitary waves in diatomic lattices include persistence results near the equal mass ratio limit [20], asymptotic approximations near the long-wave limit [21,22], and numerically assisted study of radiation generated from long-wave solitons in the time evolution [23].

We shall now present the main approximation theorem. We use the standard notation $\ell^{2}$ to denote square summable sequences equipped with the norm

$$
\|u\|_{\ell^{2}}:=\left(\sum_{k \in \mathbb{Z}}\left|u_{k}\right|^{2}\right)^{1 / 2}
$$

from which it is obvious that $\sup _{k \in \mathbb{Z}}\left|u_{k}\right| \leq\|u\|_{\ell^{2}}$. Another useful property of the $\ell^{2}$ space is being a Banach algebra with respect to pointwise multiplication.

Theorem 1. Assume that $Q^{*} \in C^{1}\left(\left[0, T_{0}\right], \ell^{2}\right)$ is a solution of the scalar FPU lattice (5) with $W \in C^{3}(\mathbb{R})$ and $W^{\prime \prime}(0)>0$ for a fixed $T_{0}>0$. There exist $\varepsilon_{0}>0, C_{0}>0$, and $C>0$ such that for all $\varepsilon \in\left(0, \varepsilon_{0}\right)$, the following is true. If $(Q(0), q(0)) \in \ell^{2} \times \ell^{2}$ satisfy the bound

$$
\begin{equation*}
\sup _{j \in 2 \mathbb{Z}}\left(\left|Q_{j}(0)-Q_{j}^{*}(0)\right|+\left|q_{j+1}(0)-\frac{Q_{j+2}^{*}(0)+Q_{j}^{*}(0)}{2}\right|\right) \leq \varepsilon \tag{6}
\end{equation*}
$$

and $Q^{*} \in C^{1}\left(\left[0, T_{0}\right], \ell^{2}\right)$ satisfy the bound

$$
\begin{equation*}
\sup _{t \in\left[0, T_{0}\right]} \sup _{j \in 2 \mathbb{Z}}\left|Q_{j+2}^{*}(t)-Q_{j}^{*}(t)\right| \leq C_{0} \tag{7}
\end{equation*}
$$

then there exists the unique solution $(Q, q) \in C^{1}\left(\left[0, T_{0}\right], \ell^{2} \times \ell^{2}\right)$ to the diatomic FPU system (2)-(3), which satisfies the bound

$$
\begin{equation*}
\sup _{t \in\left[0, T_{0}\right]} \sup _{j \in 2 \mathbb{Z}}\left(\left|Q_{j}(t)-Q_{j}^{*}(t)\right|+\left|q_{j+1}(t)-\frac{Q_{j+2}^{*}(t)+Q_{j}^{*}(t)}{2}\right|\right) \leq C \varepsilon \tag{8}
\end{equation*}
$$

Remark 2. The approximation result of Theorem 1 is nontrivial since the right hand side of the associated first order system to (2), and (3) multiplied with $\varepsilon^{-2}$, is of order $\mathcal{O}\left(\varepsilon^{-1}\right)$. Standard Gronwall's inequality only gives estimates on an $\mathcal{O}(\varepsilon)$-time scale and not on the natural $\mathcal{O}(1)$-time scale.

Remark 3. Approximation results for systems with a small perturbation parameter in front of the time derivatives, similar to system (2)-(3) have been considered in [24]. However, the abstract theorem from [24] does not apply since the nonlinear interaction appearing here is different from the one considered in Eq. (14) of [24]. The approach in [24] is based on a normal form transformation, whereas the proof presented here is based on a suitable choice of coordinates and energy estimates.

Remark 4. The monoatomic FPU system (5) is also Hamiltonian with the total energy

$$
\begin{equation*}
H_{\mathrm{FPU}}=\sum_{j \in 2 \mathbb{Z}} \frac{1}{2} \dot{Q}_{j}^{2}+2 W\left(\frac{Q_{j+2}-Q_{j}}{2}\right) \tag{9}
\end{equation*}
$$

Since $W \in C^{3}(\mathbb{R})$ and $W^{\prime \prime}(0)>0$, the conserved energy (9) is coercive for small displacements. As a result, the constraint (7) is verified for all times if the initial condition for $Q^{*} \in C^{1}\left(\left[0, T_{0}\right], \ell^{2}\right)$ yields a sufficiently small value for $H_{\mathrm{FPU}}$ due to small displacements and small velocities.

The remainder of the paper is organized as follows. In Section 2, we rewrite the diatomic FPU system in new coordinates which are more suitable to express perturbations to the motion given by the limit system (4) and (5). The bounds in Theorem 1 are obtained with the energy estimates in Section 3 for the simple case with $W^{\prime}(u)=u+u^{2}$. Generalizations to other nonlinear interaction potentials $W(u)$ are discussed in Section 4.

## 2. Change of coordinates

By using suitable chosen coordinates, we will separate the fast and slow dynamics of the diatomic FPU system (2)-(3) and will introduce perturbations to the motion given by the limit system (4)-(5). Note that the same choice of coordinates was used in [7] in the study of traveling waves. Let us set

$$
U_{j}:=\frac{1}{2}\left(Q_{j+2}-Q_{j}\right) \quad \text { and } \quad w_{j+1}:=q_{j+1}-\frac{1}{2}\left(Q_{j+2}+Q_{j}\right),
$$

so that

$$
q_{j+1}-Q_{j}=U_{j}+w_{j+1} \quad \text { and } \quad Q_{j+2}-q_{j+1}=U_{j}-w_{j+1} .
$$

The diatomic FPU system (2)-(3) is now written as

$$
2 \ddot{U}_{j}=W^{\prime}\left(U_{j+2}+w_{j+3}\right)-W^{\prime}\left(U_{j}-w_{j+1}\right)-W^{\prime}\left(U_{j}+w_{j+1}\right)+W^{\prime}\left(U_{j-2}-w_{j-1}\right)
$$

and

$$
\begin{aligned}
\varepsilon^{2} \ddot{w}_{j+1}= & W^{\prime}\left(U_{j}-w_{j+1}\right)-W^{\prime}\left(U_{j}+w_{j+1}\right)-\frac{1}{2} \varepsilon^{2} W^{\prime}\left(U_{j+2}+w_{j+3}\right) \\
& +\frac{1}{2} \varepsilon^{2} W^{\prime}\left(U_{j}-w_{j+1}\right)-\frac{1}{2} \varepsilon^{2} W^{\prime}\left(U_{j}+w_{j+1}\right)+\frac{1}{2} \varepsilon^{2} W^{\prime}\left(U_{j-2}-w_{j-1}\right) .
\end{aligned}
$$

For the particular choice $W^{\prime}(u)=u+u^{2}$, we obtain

$$
\begin{aligned}
& W^{\prime}\left(U_{j}-w_{j+1}\right)+W^{\prime}\left(U_{j}+w_{j+1}\right)=2 U_{j}+2 U_{j}^{2}+2 w_{j+1}^{2}, \\
& W^{\prime}\left(U_{j}-w_{j+1}\right)-W^{\prime}\left(U_{j}+w_{j+1}\right)=-2 w_{j+1}-4 U_{j} w_{j+1},
\end{aligned}
$$

which yields the following system of equations:

$$
\begin{align*}
\ddot{U}_{j}+U_{j}+U_{j}^{2}+w_{j+1}^{2} & =g\left(U_{j+2}, U_{j-2}, w_{j+3}, w_{j-1}\right),  \tag{10}\\
\varepsilon^{2} \ddot{w}_{j+1}+\left(2+\varepsilon^{2}\right) w_{j+1}\left(1+2 U_{j}\right) & =\varepsilon^{2} h\left(U_{j+2}, U_{j-2}, w_{j+3}, w_{j-1}\right), \tag{11}
\end{align*}
$$

where

$$
\begin{align*}
& g\left(U_{j+2}, U_{j-2}, w_{j+3}, w_{j-1}\right)=\frac{1}{2} W^{\prime}\left(U_{j+2}+w_{j+3}\right)+\frac{1}{2} W^{\prime}\left(U_{j-2}-w_{j-1}\right),  \tag{12}\\
& h\left(U_{j+2}, U_{j-2}, w_{j+3}, w_{j-1}\right)=-\frac{1}{2} W^{\prime}\left(U_{j+2}+w_{j+3}\right)+\frac{1}{2} W^{\prime}\left(U_{j-2}-w_{j-1}\right) . \tag{13}
\end{align*}
$$

The dynamics of $U$ and $w$ occurs now at two different scales: $U$ changes on the time scale of $\mathcal{O}(1)$, whereas $w$ changes on the faster time scale of $\mathcal{O}(\varepsilon)$. The approximation result of Theorem 1 justifies the dynamics of $U$ on the time scale of $\mathcal{O}(1)$. The dynamics of $w$ is slaved to the dynamics of $U$ on this time scale.

## 3. The error estimates

The leading-order approximation in the new coordinates is denoted by $(U, w)=(\Psi, 0)$, where $\Psi$ satisfies

$$
\begin{equation*}
\ddot{\Psi}_{j}+\Psi_{j}+\Psi_{j}^{2}=g\left(\Psi_{j+2}, \Psi_{j-2}, 0,0\right) . \tag{14}
\end{equation*}
$$

After inserting this approximation into the equations of motion (10) and (11), the remaining terms are collected in the residual, which is given by

$$
\begin{aligned}
\operatorname{Res}_{U, j} & =0 \\
\operatorname{Res}_{w, j} & =\varepsilon^{2} h\left(\Psi_{j+2}, \Psi_{j-2}, 0,0\right)
\end{aligned}
$$

The residual terms obey the following estimate.
Lemma 5. Assume that $\Psi \in C\left(\left[0, T_{0}\right], \ell^{2}\right)$ is a solution of the scalar equation (14) for some $T_{0}>0$. Then there exists a constant $C>0$ that depends on $\Psi$ such that for all $\varepsilon \in(0,1)$ we have

$$
\begin{equation*}
\sup _{t \in\left[0, T_{0}\right]}\left\|R e s_{w}\right\|_{\ell^{2}} \leq C \varepsilon^{2} \tag{15}
\end{equation*}
$$

Proof. We recall that $\ell^{2}$ is a Banach algebra with respect to pointwise multiplication. Due to this property, it follows from (13) with $W^{\prime}(u)=u+u^{2}$ that

$$
\left\|\operatorname{Res}_{w}\right\|_{\ell^{2}} \leq \varepsilon^{2}\left(\|\Psi\|_{\ell^{2}}+\|\Psi\|_{\ell^{2}}^{2}\right),
$$

which gives (15) under the condition $\Psi \in C\left(\left[0, T_{0}\right], \ell^{2}\right)$.
For estimating the difference between the approximation and the true solution we introduce the error functions $R$ and $v$ by using the decomposition

$$
\begin{equation*}
U_{j}=\Psi_{j}+\varepsilon R_{j} \quad \text { and } \quad w_{j+1}=\varepsilon v_{j+1} \tag{16}
\end{equation*}
$$

These functions satisfy the following system

$$
\begin{align*}
& \ddot{R}_{j}+R_{j}+2 \Psi_{j} R_{j}+\varepsilon R_{j}^{2}+\varepsilon v_{j+1}^{2}=L_{U, j}(\Psi)(R, v)+\varepsilon N_{U, j}(\Psi, R, v),  \tag{17}\\
& \varepsilon^{2} \ddot{v}_{j+1}+2 v_{j+1}\left(1+2 \Psi_{j}+2 \varepsilon R_{j}\right)=\varepsilon^{2} L_{w, j}(\Psi)(R, v)+\varepsilon^{3} N_{w, j}(\Psi, R, v)+\varepsilon^{-1} \operatorname{Res}_{w, j} \tag{18}
\end{align*}
$$

where the linear terms in $(R, v)$ are given by

$$
\begin{align*}
L_{U, j}(\Psi)(R, v)= & \frac{1}{2}\left(R_{j+2}+R_{j-2}\right)+\frac{1}{2}\left(v_{j+3}-v_{j-1}\right)  \tag{19}\\
& +\Psi_{j+2}\left(R_{j+2}+v_{j+3}\right)+\Psi_{j-2}\left(R_{j-2}-v_{j-1}\right), \\
L_{w, j}(\Psi)(R, v)= & -\left(1+2 \Psi_{j}\right) v_{j+1}-\frac{1}{2}\left(R_{j+2}-R_{j-2}\right)-\frac{1}{2}\left(v_{j+3}+v_{j-1}\right)  \tag{20}\\
& -\Psi_{j+2}\left(R_{j+2}+v_{j+3}\right)+\Psi_{j-2}\left(R_{j-2}-v_{j-1}\right),
\end{align*}
$$

and quadratic terms in $(R, v)$ are given by

$$
\begin{align*}
& N_{U, j}(\Psi, R, v)=\frac{1}{2}\left(R_{j+2}+v_{j+3}\right)^{2}+\frac{1}{2}\left(R_{j-2}-v_{j-1}\right)^{2}  \tag{21}\\
& N_{w, j}(\Psi, R, v)=-2 R_{j} v_{j+1}-\frac{1}{2}\left(R_{j+2}+v_{j+3}\right)^{2}+\frac{1}{2}\left(R_{j-2}-v_{j-1}\right)^{2} . \tag{22}
\end{align*}
$$

The linear and quadratic terms obey the following estimate.
Lemma 6. Assume that $\Psi \in C\left(\left[0, T_{0}\right], \ell^{2}\right)$ is a solution of the scalar equation (14) for some $T_{0}>0$. Then there exists a constant $C>0$ that depends on $\Psi$ such that for all $\varepsilon \in(0,1)$ we have

$$
\begin{align*}
\left\|L_{U}(\Psi)(R, v)\right\|_{\ell^{2}}+\left\|L_{w}(\Psi)(R, v)\right\|_{\ell^{2}} & \leq C\left(\|R\|_{\ell^{2}}+\|v\|_{\ell^{2}}\right)  \tag{23}\\
\left\|N_{U}(\Psi, R, v)\right\|_{\ell^{2}}+\left\|N_{w}(\Psi, R, v)\right\|_{\ell^{2}} & \leq C\left(\|R\|_{\ell^{2}}^{2}+\|v\|_{\ell^{2}}^{2}\right) . \tag{24}
\end{align*}
$$

Proof. The proof follows from (19), (20), (21), and (22) due to the same property of $\ell^{2}$ being a Banach algebra with respect to pointwise multiplication.

The dynamics of the error functions is estimated with the help of a suitable chosen energy. We define the energy function by

$$
\begin{equation*}
E(t)=\frac{1}{2} \sum_{j \in 2 \mathbb{Z}} \dot{R}_{j}^{2}+R_{j}^{2}+\varepsilon^{2} \dot{v}_{j+1}^{2}+2 v_{j+1}^{2}+2 \Psi_{j}\left(R_{j}^{2}+2 v_{j+1}^{2}\right)+4 \varepsilon R_{j} v_{j+1}^{2} \tag{25}
\end{equation*}
$$

Computing the time derivative of $E(t)$ yields

$$
\begin{align*}
\frac{d}{d t} E(t)= & \left\langle\dot{R}, \ddot{R}+R+2 \Psi R+2 \varepsilon v^{2}\right\rangle_{\ell^{2}}  \tag{26}\\
& +\left\langle\dot{v}, \varepsilon^{2} \ddot{v}+2 v+4 \Psi v+4 \varepsilon R v\right\rangle_{\ell^{2}}+\left\langle\dot{\Psi}, R^{2}+2 v^{2}\right\rangle_{\ell^{2}}
\end{align*}
$$

where $(\Psi R)_{j}=\Psi_{j} R_{j}$ and $(\Psi v)_{j}=\Psi_{j} v_{j+1}$. By substituting the dynamical equations for ( $R, v$ ) into (26), we obtain

$$
\begin{align*}
\frac{d}{d t} E(t)= & \left\langle\dot{R},-\varepsilon R^{2}+\varepsilon v^{2}+L_{U}(\Psi)(R, v)+\varepsilon N_{U}(\Psi, R, v)\right\rangle_{\ell^{2}}  \tag{27}\\
& +\left\langle\varepsilon \dot{v}, \varepsilon L_{w}(\Psi)(R, v)+\varepsilon^{2} N_{w}(\Psi, R, v)+\varepsilon^{-2} \operatorname{Res}_{w}\right\rangle_{\ell^{2}} \\
& +\left\langle\dot{\Psi}, R^{2}+2 v^{2}\right\rangle_{\ell^{2}} .
\end{align*}
$$

The energy function controls the perturbations and their time derivative if displacements of heavy particles relatively to each other are sufficiently small, as in the condition (7) of Theorem 1. The following lemma gives the corresponding result.

Lemma 7. Assume that $\Psi \in C\left(\left[0, T_{0}\right], \ell^{2}\right)$ is a solution of the scalar equation (14) for some $T_{0}>0$. There exists $C>0$ such that if

$$
\begin{equation*}
C_{0}:=\sup _{t \in\left[0, T_{0}\right]} \sup _{j \in 2 \mathbb{Z}}\left|\Psi_{j}(t)\right|<\frac{1}{4}, \tag{28}
\end{equation*}
$$

and $\varepsilon\|R(t)\|_{\ell^{2}} \leq \frac{1}{4}$ then

$$
\begin{equation*}
\|\dot{R}(t)\|_{\ell^{2}}^{2}+\|R(t)\|_{\ell^{2}}^{2}+\|\varepsilon \dot{v}(t)\|_{\ell^{2}}^{2}+\|v(t)\|_{\ell^{2}}^{2} \leq C E(t) . \tag{29}
\end{equation*}
$$

Proof. We obtain from (25) by using the bound (28) that

$$
2 E(t) \geq\|\dot{R}(t)\|_{\ell^{2}}^{2}+\left(1-2 C_{0}\right)\|R(t)\|_{\ell^{2}}^{2}+\|\varepsilon \dot{v}(t)\|_{\ell^{2}}^{2}+2\left(1-2 C_{0}\right)\|v(t)\|_{\ell^{2}}^{2}-4 \varepsilon\|R(t)\|_{\ell^{2}}\|v(t)\|_{\ell^{2}}^{2},
$$

which yields (29) for some constant $C>0$ provided that $C_{0}<\frac{1}{4}$ and $\varepsilon\|R(t)\|_{\ell^{2}} \leq \frac{1}{4}$.
The essential point in the proof of the approximation result of Theorem 1 is that the fast dynamics of $v$ can be controlled by the $\|\varepsilon \dot{v}(t)\|_{\ell^{2}}^{2}$ term in the energy bound (29) and in the energy balance equation (27). The following lemma gives the useful estimate from the energy balance equation (27).

Lemma 8. Assume that $\Psi \in C^{1}\left(\left[0, T_{0}\right], \ell^{2}\right)$ is a solution of the scalar equation (14) for some $T_{0}>0$ satisfying (28). There exist constants $C_{1}, C_{2}, C_{3}>0$ that depend on $\Psi$ such that for all $\varepsilon \in(0,1)$ we have

$$
\begin{equation*}
\frac{d}{d t} E(t) \leq C_{1} E(t)^{1 / 2}+C_{2} E(t)+C_{3} \varepsilon E(t)^{3 / 2}, \quad t \in\left[0, T_{0}\right] \tag{30}
\end{equation*}
$$

as long as $\varepsilon\|R(t)\|_{\ell^{2}} \leq \frac{1}{4}$.
Proof. We use the Cauchy-Schwarz inequality in (27) together with the estimates (15), (23), (24), and (29). This yields (30).

We can now conclude the proof of Theorem 1. Let $S(t):=E(t)^{1 / 2}$. The initial bound (6) yields $S(0) \leq C_{0}$ for some $C_{0}>0$ independently of $\varepsilon \in\left(0, \varepsilon_{0}\right)$. The energy balance estimate (30) can be rewritten in the form

$$
\begin{equation*}
\frac{d}{d t} S(t) \leq C_{1}+C_{2} S(t)+C_{3} \varepsilon S(t)^{2}, \quad t \in\left[0, T_{0}\right] \tag{3}
\end{equation*}
$$

where the constants $C_{1}, C_{2}, C_{3}>0$ have been redefined. Let $T_{*}$ be defined by

$$
T_{*}:=\sup \left\{T>0: \quad \varepsilon S(t) \leq \frac{C_{2}}{C_{3}}, \quad \varepsilon\|R(t)\|_{\ell^{2}} \leq \frac{1}{4}, \quad t \in[0, T]\right\}
$$

for the given constants $\varepsilon, C_{2}$, and $C_{3}$. Then, by Gronwall's inequality, we obtain

$$
S(t) \leq\left[S(0)+\left(2 C_{2}\right)^{-1} C_{1}\right] e^{2 C_{2} t} \leq\left[C_{0}+\left(2 C_{2}\right)^{-1} C_{1}\right] e^{2 C_{2} T_{0}}, \quad t \in\left[0, T_{0}\right] .
$$

Since $T_{0}<T_{*}$ if $\varepsilon>0$ is appropriately small, we obtain $S(t) \leq C$ for some $C>0$ independently of $\varepsilon \in\left(0, \varepsilon_{0}\right)$, and the final bound (8) holds. It also follows from the energy bound (29) that $\|R(t)\|_{\ell^{2}} \leq C$ for some $C>0$. The approximation result of Theorem 1 is proven.

## 4. Generalization

We have proven the approximation result of Theorem 1 for the simplest nonlinear interaction potential $W^{\prime}(u)=u+u^{2}$. For a more general interaction potential $W \in C^{3}(\mathbb{R})$, Taylor expansions around $U$ yield

$$
W^{\prime}\left(U_{j}-w_{j+1}\right)+W^{\prime}\left(U_{j}+w_{j+1}\right)=2 W^{\prime}\left(U_{j}\right)+\mathcal{O}\left(\left|w_{j+1}\right|^{2}\right)
$$

and

$$
W^{\prime}\left(U_{j}-w_{j+1}\right)-W^{\prime}\left(U_{j}+w_{j+1}\right)=-2 W^{\prime \prime}\left(U_{j}\right) w_{j+1}+\mathcal{O}\left(\left|w_{j+1}\right|^{2}\right)
$$

so that the system of coupled equations (10) and (11) is rewritten in the more general form:

$$
\begin{align*}
\ddot{U}_{j}+W^{\prime}\left(U_{j}\right)+\mathcal{O}\left(\left|w_{j+1}\right|^{2}\right) & =g\left(U_{j+2}, U_{j-2}, w_{j+3}, w_{j-1}\right),  \tag{32}\\
\varepsilon^{2} \ddot{w}_{j+1}+\left(2+\varepsilon^{2}\right) W^{\prime \prime}\left(U_{j}\right) w_{j+1}+\mathcal{O}\left(\left|w_{j+1}\right|^{2}\right) & =\varepsilon^{2} h\left(U_{j+2}, U_{j-2}, w_{j+3}, w_{j-1}\right) \tag{33}
\end{align*}
$$

The energy function for the perturbation terms in the decomposition (16) becomes

$$
\begin{equation*}
E(t)=\frac{1}{2} \sum_{j \in 2 \mathbb{Z}} \dot{R}_{j}^{2}+W^{\prime \prime}\left(\Psi_{j}\right) R_{j}^{2}+\varepsilon^{2} \dot{v}_{j+1}^{2}+2 W^{\prime \prime}\left(\Psi_{j}+\varepsilon R_{j}\right) v_{j+1}^{2} . \tag{34}
\end{equation*}
$$

It follows by repeating the previous analysis that the same approximation result stated in Theorem 1 applies to the more general interaction potential satisfying the conditions $W \in C^{3}(\mathbb{R})$ and $W^{\prime \prime}(0)>0$.

## Acknowledgments

G. Schneider is partially supported by the Deutsche Forschungsgemeinschaft DFG through the CRC 1173 "Wave phenomena". D.E. Pelinovsky is partially supported by the grant of the President of the Russian Federation for leading scientific schools of the Russian Federation NSh-2485.2020.5.

## References

[1] M.A. Collins, Solitons in the diatomic chain, Phys. Rev. A 31 (3) (1985) 1754-1762.
[2] G. Huang, Soliton excitations in one-dimensional diatomic lattices, Phys. Rev. B 51 (18) (1995) 12347-12360.
[3] Y. Tabata, Stable solitary wave in diatomic Toda lattice, J. Phys. Soc. Japan 65 (12) (1996) 3689-3691.
[4] K.R. Jayaprakash, Yu. Starosvetsky, A.F. Vakakis, New family of solitary waves in granular dimer chains with no precompression, Phys. Rev. E 83 (2011) 11, 036606.
[5] K.R. Jayaprakash, Yu. Starosvetsky, A.F. Vakakis, O.V. Gendelman, Nonlinear resonances leading to strong pulse attenuation in granular dimer chains, J. Nonlinear Sci. 23 (3) (2013) 363-392.
[6] M.A. Porter, C. Daraio, I. Szelengowics, E.B. Herbold, P.G. Kevrekidis, Highly nonlinear solitary waves in heterogeneous periodic granular media, Physica D 238 (2009) 666-676.
[7] A. Hoffman, J.D. Wright, Nanopteron solutions of diatomic Fermi-Pasta-Ulam-Tsingou lattices with small mass-ratio, Physica D 358 (2017) 33-59.
[8] T.E. Faver, J.D. Wright, Exact diatomic Fermi-Pasta-Ulam-Tsingou solitary waves with optical band ripples at infinity, SIAM J. Math. Anal. 50 (2018) 182-250.
[9] C.J. Lustri, M.A. Porter, Nanoptera in a period-2 Toda chain, SIAM J. Appl. Dyn. Syst. 17 (2018) 1182-1212.
[10] C.J. Lustri, Nanoptera and Stokes curves in the 2-periodic Fermi-Pasta-Ulam-Tsingou equation, Physica D 402 (2020) 13, 132239.
[11] A. Vainchtein, Y. Starosvetsky, J.D. Wright, R. Perline, Solitary waves in diatomic chains, Phys. Rev. E 93 (4) (2016) 042210.
[12] R. Livi, M. Spicci, R.S. MacKay, Breathers on a diatomic FPU chain, Nonlinearity 10 (6) (1997) 1421-1434.
[13] G. James, P. Noble, Breathers on diatomic Fermi-Pasta-Ulam lattices, Physica D 196 (1-2) (2004) 124-171.
[14] G. James, P. Noble, Weak coupling limit and localized oscillations in Euclidean invariant Hamiltonian systems, J. Nonlinear Sci. 18 (2008) 433-461.
[15] K. Yoshimura, Existence and stability of discrete breathers in diatomic Fermi-Pasta-Ulam type lattices, Nonlinearity 24 (4) (2011) 293-317.
[16] G. James, Periodic travelling waves and compactons in granular chains, J. Nonlinear Sci. 22 (2012) 813-848.
[17] M. Betti, D.E. Pelinovsky, Periodic travelling waves in diatomic granular chains, J. Nonlinear Sci. 23 (2013) 689-730.
[18] W.X. Qin, Modulation of uniform motion in diatomic Frenkel-Kontorova model, Discrete Contin. Dyn. Syst. 34 (9) (2014) 3773-3788.
[19] W.X. Qin, Wave propagation in diatomic lattices, SIAM J. Math. Anal. 47 (1) (2015) 477-497.
[20] T.E. Faver, H.J. Hupkes, Micropteron traveling waves in diatomic Fermi-Pasta-Ulam-Tsingou lattices under the equal mass limit, Physica D (2020).
[21] J. Gaison, S. Moskow, J.D. Wright, Q. Zhang, Approximation of polyatomic FPU lattices by KdV equations, Multiscale Model. Simul. 12 (2014) 953-995.
[22] J.A.D. Wattis, Asymptotic approximations to travelling waves in the diatomic Fermi-Pasta-Ulam lattice, Math. Eng. 1 (2019) 327-342.
[23] N. Giardetti, A. Shapiro, S. Windle, J.D. Wright, Metastability of solitary waves in diatomic FPUT lattices, Math. Eng. 1 (2019) 419-433.
[24] S. Baumstark, G. Schneider, K. Schratz, D. Zimmermann, Effective slow dynamics models for a class of dispersive systems, J. Dynam. Differential Equations (2019).


[^0]:    * Corresponding author at: Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada, L8S 4K1.

    E-mail addresses: dmpeli@math.mcmaster.ca (D.E. Pelinovsky), guido.schneider@mathematik.uni-stuttgart.de
    (G. Schneider).

