HAL
open science

Bogolyubov's averaging theorem applied to the Kramers-Henneberger Hamiltonian

Elena Floriani, Jonathan Dubois, Cristel Chandre

To cite this version:

Elena Floriani, Jonathan Dubois, Cristel Chandre. Bogolyubov's averaging theorem applied to the Kramers-Henneberger Hamiltonian. Physica D: Nonlinear Phenomena, 2022, 431, pp.133124. 10.1016/j.physd.2021.133124 . hal-03277919

HAL Id: hal-03277919

https://hal.science/hal-03277919v1

Submitted on 5 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Bogolyubov's averaging theorem applied to the Kramers-Henneberger Hamiltonian

E. Floriani[a], J. Dubois[b], C. Chandre[c]

[a] *Aix-Marseille Univ, Université de Toulon, CNRS, CPT, Marseille, France*
[b] *Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Straße 38, Dresden, 01187, Germany*
[c] *CNRS, Aix-Marseille Université, Centrale Marseille, I2M, Marseille, France*

Abstract

We apply Bogolyubov's averaging theorem to the motion of an electron of an atom driven by a linearly polarized laser field in the Kramers-Henneberger frame. We provide estimates of the differences between the original trajectories and the trajectories associated with the averaged system as a function of the parameters of the laser field and the region of phase space. We formulate a modified Bogolyubov averaging theorem based on the Hamiltonian properties of the system, and show that this version is better suited for these systems. From these estimates, we discuss the validity of the Kramers-Henneberger approximation.

Keywords: Averaging theorem, Hamiltonian systems, Kramers-Henneberger approximation

1. Introduction

The motion of electrons in a combined Coulomb and time-dependent electric field displays a rather wide variety of behaviours which are easily understood only in two limits: First, when the electron is close to its ionic core, its motion is mainly driven by the Coulomb force, which corresponds to a case where the electric field acts as a small perturbation of the integrable case, therefore it evolves typically on invariant tori. Second, when the electron is far away from the ionic core, the electron is mostly driven by the electric field and it ionizes quickly. This second limit also corresponds to an integrable situation or a small perturbation of it. In between these two limits, the motion

of the electron cannot be treated perturbatively, and typically displays features at different spatial and temporal scales. Remarkably, this is the most interesting situation dynamically and also for practical applications. Indeed, high harmonics of the driving field are generated by these trajectories which wander chaotically between the two integrable situations [1]. For example, the so-called recollisions, the keystone of strong-field atomic physics [2, 3, 4], return to the ionic core after a pre-ionization. Another example is afforded by the ionization stabilization in strong fields [5, 6, 7] where the ionization rate decreases as intensity is increased and a stable state is created as a result of the combination of the Coulomb potential and the time-dependent electric field.

A natural way to understand some of the dynamical features experienced by the electron in a combined Coulomb and electric field is to perform averaging over the smaller temporal scales (see, e.g., Refs. [24, 25]). It allows the reduction of phase space by focusing on more relevant time scales. An important milestone [8, 9] was achieved by Kramers and Henneberger with a change of coordinates adapted to the electric field and an averaging of the potential in these new coordinates. In this frame, the potential highlights a local minimum and therefore the possibility for a stable state, dressed by the field, far away from the ionic core. This has given rise to a quite extensive literature on the Kramers-Henneberger (KH) approximation [10, 11, 12, 13, 14, 15, 16] which led to the denomination of Kramers-Henneberger atom [17, 18, 19]. The importance of this reduction is undeniable. The main question we address here is how valid the approximation is, and if so, under which conditions. This question has been addressed in Refs. [13, 20]. Classically, it amounts to estimating the distance between the trajectories of the original system with the ones of the averaged system. The averaging would be valid if this distance is small enough for at least the duration of the laser pulse (at least of the order of a few laser cycles). We follow Ref. [20] to derive rigorous estimates between the trajectories using Bogolyubov's averaging theorem. Bogolyubov's theorem is used to determine in which regions of phase space and for which region in parameter space an averaging method is valid for a given time. Here we show that this theorem, as stated in Refs. [21, 22, 23] cannot provide sufficiently accurate analytic estimates to fully address the question of the validity. We illustrate this point using a soft-Coulomb potential, which fully takes into account the long-range Coulomb interaction.

In order to provide estimates closer to the ionic core region, we formulate a modification of Bogolyubov's averaging theorem which is more adapted

to Hamiltonian systems. The resulting estimates are shown to be much more accurate than the ones of the original formulation of Bogolyubov's theorem. They allow us to bring more elements on the validity of the KH approximation. This question cannot be simply answered by a consideration of the values of the parameters of the laser field only: the regions in phase space which are visited by the trajectories are paramount.

Here, we consider the Hamiltonian of an electron in a combined ion potential and electric field in the dipole approximation

$$H_{\mathrm{e}}(\mathbf{r}_{\mathrm{e}}, \mathbf{p}_{\mathrm{e}}, t) = \frac{\mathbf{p}_{\mathrm{e}}^2}{2} + V(\mathbf{r}_{\mathrm{e}}) + \mathbf{r}_{\mathrm{e}} \cdot \mathbf{E}(t)\,, \tag{1}$$

where $\mathbf{r}_{\mathrm{e}}$ is the position of the electron and $\mathbf{p}_{\mathrm{e}}$ its canonically conjugate momentum. The electron-ion potential $V(\mathbf{r}_{\mathrm{e}}) = -(|\mathbf{r}_{\mathrm{e}}|^2 + a^2)^{-1/2}$ is modeled as a soft-Coulomb potential. The electric field is linearly polarized and is given by

$$\mathbf{E}(t) = E_0 \cos(\omega t)\, \hat{\mathbf{x}} = -\frac{\partial \mathbf{A}}{\partial t}\,,$$

where $\mathbf{A}(t)$ is the vector potential.

We move to the KH frame by performing the canonical change of coordinates

$$\mathbf{p} = \mathbf{p}_{\mathrm{e}} - \mathbf{A}(t)\,, \tag{2a}$$

$$\mathbf{r} = \mathbf{r}_{\mathrm{e}} - \frac{\mathbf{E}(t)}{\omega^2}\,. \tag{2b}$$

In the KH frame, Hamiltonian (1) becomes

$$H(\mathbf{r}, \mathbf{p}, t) = \frac{\mathbf{p}^2}{2} + V\left(\mathbf{r} + \frac{\mathbf{E}(t)}{\omega^2}\right)\,. \tag{3}$$

In the literature on KH frame, an approximation -coined Kramers-Henneberger approximation- is performed in addition to the change of coordinates. In this approximation, the above time-dependent Hamiltonian is averaged over one period of the laser field, and becomes

$$\langle H \rangle(\mathbf{r}, \mathbf{p}) = \frac{\mathbf{p}^2}{2} + V_{\mathrm{KH}}(\mathbf{r})\,, \tag{4}$$

where

$$V_{\mathrm{KH}}(\mathbf{r}) = \frac{1}{T} \int_0^T V\left(\mathbf{r} + \frac{\mathbf{E}(t)}{\omega^2}\right)\, \mathrm{d}t\,.$$

The main gain is that $\langle H \rangle$ is now time-independent, and it now corresponds to a conserved quantity. As a consequence, its dynamics is more easily understood. However, a loss of dynamical information is associated with this approximation, since $\langle H \rangle$ is no longer conjugated to H. The central question is how much information has been lost, moving from H to $\langle H \rangle$, and overall, how good the KH approximation is. In order to bring elements to answer this question, we compare the trajectories of Hamiltonian (3) with the ones of Hamiltonian (4). Equivalently, this brings some comparison between the trajectories $(\mathbf{r}_{\mathrm{e}}(t), \mathbf{p}_{\mathrm{e}}(t))$ of the original system (1) and the reconstructed trajectories $(\overline{\mathbf{r}}(t) + \mathbf{E}(t)/\omega^2, \overline{\mathbf{p}}(t) + \mathbf{A}(t))$ where $(\overline{\mathbf{r}}(t), \overline{\mathbf{p}}(t))$ are trajectories of Hamiltonian (4).

In Sec. 2, we recall the formulation of Bogolyubov's averaging theorem and provide a concise proof. We provide and discuss the resulting estimates for a soft-Coulomb potential in one dimension. Next, in Sec. 3, we modify Bogolyubov's theorem by adapting the estimates to a Hamiltonian flow. We compare and discuss the resulting estimates on a one-dimensional soft-Coulomb potential.

2. Bogolyubov's averaging theorem

2.1. Statement of the theorem

We consider Bogolyubov's averaging theorem for comparing the trajectories of the original system with the ones of averaged system. This was the approach followed in Ref. [20] to assess the validity of the Kramers-Henneberger approximation. We briefly recall the statement of the theorem and a brief proof for the periodic case (see Refs. [21, 22, 23] for more details). We consider the following equations of motion:

$$\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}t} = \varepsilon\, \mathbf{Z}(t, \mathbf{z}, \varepsilon)\,,$$

where ε is a small positive parameter and $\mathbf{z} \in \mathbb{R}^n$ are the coordinates of the electron in phase space. Two Cauchy problems are compared:

$$\mathrm{I}: \quad \frac{\mathrm{d}\mathbf{z}}{\mathrm{d}t} = \varepsilon\, \mathbf{Z}(t, \mathbf{z}, \varepsilon)\,, \qquad \text{with } \mathbf{z}(0) = \mathbf{z}_0\,, \tag{5a}$$

$$\mathrm{II}: \quad \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = \varepsilon\, \langle \mathbf{Z} \rangle(\mathbf{u}, \varepsilon)\,, \qquad \text{with } \mathbf{u}(0) = \mathbf{z}_0\,. \tag{5b}$$

Here $\mathbf{Z}(t, \mathbf{z}, \varepsilon)$ is defined in a domain $\mathcal{D}$ of $\mathbf{z}$. For $t \geq 0$ and $0 \leq \varepsilon \leq \varepsilon_1$ the following conditions are assumed:

(a) $\mathbf{Z}(t, \mathbf{z}, \varepsilon)$ is a t-measurable function for fixed $\mathbf{z}$ and ε, and the limit

$$\langle \mathbf{Z} \rangle(\mathbf{z}, \varepsilon) = \lim_{T \to \infty} \frac{1}{T} \int_0^T \mathbf{Z}(t, \mathbf{z}, \varepsilon) \, \mathrm{d}t \,, \tag{6}$$

exists uniformly relative to both $\mathbf{z} \in \mathcal{D}$ and $\varepsilon \in [0, \varepsilon_1]$.

(b) It is possible to find constants M and λ such that, for any $\mathbf{z}, \mathbf{z}'$ in $\mathcal{D}$,

$$\|\mathbf{Z}(\mathbf{z}, \varepsilon)\| \leq M \,, \tag{7a}$$

$$\|\mathbf{Z}(t, \mathbf{z}, \varepsilon) - \mathbf{Z}(t, \mathbf{z}', \varepsilon)\| \leq \lambda \|\mathbf{z} - \mathbf{z}'\| \,. \tag{7b}$$

We consider the case where $\mathbf{Z}(t, \mathbf{z}, \varepsilon)$ is periodic in t, i.e., there exists T such that $\mathbf{Z}(t, \mathbf{z}, \varepsilon) = \mathbf{Z}(t + T, \mathbf{z}, \varepsilon)$ for all $t \geq 0$ and $\mathbf{z} \in \mathcal{D}$.

Theorem 1 ([21, 22]). *If the solution $\mathbf{u}(t)$ to the initial Cauchy problem for system II is defined for $t \geq 0$ and belongs to the domain $\mathcal{D}$, then, under Hypotheses (a) and (b), for any $L > 0$, the following inequality holds*

$$\|\mathbf{z}(t) - \mathbf{u}(t)\| \leq \varepsilon \, \mathrm{e}^{\lambda L} \, T \, (L \lambda M + 2M) \,, \tag{8}$$

for all $0 \leq t < L/\varepsilon$ and $0 \leq \varepsilon < \varepsilon_1$.

Proof: From the obvious statement,

$$\begin{aligned} \mathbf{z}(t) - \mathbf{u}(t) &= \varepsilon \int_0^t [\mathbf{Z}(\tau, \mathbf{z}(\tau), \varepsilon) - \langle \mathbf{Z} \rangle(\mathbf{u}(\tau), \varepsilon)\rangle] \mathrm{d}\tau, \\ &= \varepsilon \int_0^t [\mathbf{Z}(\tau, \mathbf{z}(\tau), \varepsilon) - \mathbf{Z}(\tau, \mathbf{u}(\tau), \varepsilon)] \, \mathrm{d}\tau \\ &\qquad + \varepsilon \int_0^t [\mathbf{Z}(\tau, \mathbf{u}(\tau), \varepsilon) - \langle \mathbf{Z} \rangle(\mathbf{u}(\tau), \varepsilon)\rangle] \, \mathrm{d}\tau, \end{aligned}$$

where we have used the fact that $\mathbf{z}(0) - \mathbf{u}(0) = \mathbf{0}$, and from Eq. (7), we deduce

$$\|\mathbf{z}(t) - \mathbf{u}(t)\| \leq \varepsilon \lambda \int_0^t \|\mathbf{z}(\tau) - \mathbf{u}(\tau)\| \, \mathrm{d}\tau + \sup_{t \in [0, L/\varepsilon]} \left\| \int_0^t \varepsilon \, \Delta \mathbf{Z}(\tau, \mathbf{u}(\tau), \varepsilon) \, \mathrm{d}\tau \right\| \,,$$

where

$$\Delta \mathbf{Z}(t, \mathbf{u}(t), \varepsilon) = \mathbf{Z}(t, \mathbf{u}(t), \varepsilon) - \langle \mathbf{Z} \rangle(\mathbf{u}(t), \varepsilon) \,.$$

Grönwall's lemma implies

$$\|\mathbf{z}(t) - \mathbf{u}(t)\| \leq \mathrm{e}^{\varepsilon\lambda t} \sup_{t\in[0,L/\varepsilon]} \left\| \int_0^t \varepsilon\, \Delta\mathbf{Z}(\tau, \mathbf{u}(\tau), \varepsilon)\, \mathrm{d}\tau \right\| .$$

Therefore, the estimate on the error $\|\mathbf{z}(t) - \mathbf{u}(t)\|$ mainly depends on the maximum gradient of the flow, i.e., λ, and on the time integral of the fluctuating part of the flow computed at the solution of the averaged equations. In particular, the exponential term does not depend on the fact that the same initial conditions were used for both systems I and II.

In the periodic case, in order to estimate the integral of $\Delta\mathbf{Z}$, the interval $[0, t]$ is partitioned: For $nT \leq t < (n+1)T$, i.e., $n = \lfloor t/T \rfloor$,

$$\left\| \int_0^t \Delta\mathbf{Z}(\tau, \mathbf{u}(\tau), \varepsilon)\, \mathrm{d}\tau \right\| \leq \sum_{i=1}^{n} \left\| \int_{(i-1)T}^{iT} \Delta\mathbf{Z}(\tau, \mathbf{u}(\tau), \varepsilon)\, \mathrm{d}\tau \right\| + \left\| \int_{nT}^{t} \Delta\mathbf{Z}(\tau, \mathbf{u}(\tau), \varepsilon)\, \mathrm{d}\tau \right\| .$$

We denote the values of $\mathbf{u}$ at each period as $\mathbf{u}_i = \mathbf{u}((i-1)T)\,,\ i = 1, \ldots, n$. Given that

$$\int_{(i-1)T}^{iT} \Delta\mathbf{Z}(\tau, \mathbf{u}_i, \varepsilon)\, \mathrm{d}\tau = 0\,,$$

which is obtained from the periodicity of the flow, we substitute this integral into the sum:

$$\begin{aligned} \left\| \int_0^t \Delta\mathbf{Z}(\tau, \mathbf{u}(\tau), \varepsilon)\, \mathrm{d}\tau \right\| &\leq \sum_{i=1}^{n} \int_{(i-1)T}^{iT} \|\Delta\mathbf{Z}(\tau, \mathbf{u}(\tau), \varepsilon) - \Delta\mathbf{Z}(\tau, \mathbf{u}_i, \varepsilon)\| \mathrm{d}\tau \\ &\qquad + \int_{nT}^{t} \|\Delta\mathbf{Z}(\tau, \mathbf{u}(\tau), \varepsilon)\|\, \mathrm{d}\tau, \\ &\leq \sum_{i=1}^{n} 2\lambda \int_{(i-1)T}^{iT} \|\mathbf{u}(\tau) - \mathbf{u}_i\|\, \mathrm{d}\tau + 2MT\,, \end{aligned} \tag{9}$$

where the inequality $\|\langle\mathbf{Z}\rangle\| \leq M$ resulting from Eqs. (6) and (7b) was used. Now, we use the fact that for $\tau \in [(i-1)T, iT]$:

$$\mathbf{u}(\tau) = \mathbf{u}_i + \int_{(i-1)T}^{\tau} \varepsilon\, \langle\mathbf{Z}\rangle(\mathbf{u}(s), \varepsilon)\, \mathrm{d}s\,,$$

so that

$$\|\mathbf{u}(\tau) - \mathbf{u}_i\| \leq \varepsilon M(\tau - (i-1)T)\,, \tag{10}$$

which in turn, implies that

$$\int_{(i-1)T}^{iT} \|\mathbf{u}(\tau) - \mathbf{u}_i\| \, \mathrm{d}\tau \leq \varepsilon M \frac{T^2}{2} \, .$$

Finally,

$$\left\| \int_0^t \Delta\mathbf{Z}(\tau, \mathbf{u}(\tau), \varepsilon) \mathrm{d}\tau \right\| \leq 2\lambda n \, \varepsilon M \, \frac{T^2}{2} + 2MT \leq T \left(\varepsilon t \lambda M + 2M \right) \, .$$

Since $\varepsilon t < L$, Eq. (8) follows.

Equation (8) is the original formulation of Bogolyubov's averaging theorem. However, when examining the proof, it can be seen that a sharper version of it can be given, essentially by introducing specific bounds for the flows $\langle \mathbf{Z} \rangle$ and $\Delta\mathbf{Z}$. We show below that this formulation improves the estimates of the discrepancy between the original and the averaged trajectories, and in fact, is better suited to a system with energy conservation. The modified statement reads as follows:

Proposition 1. *If there are constants $\overline{M}$, λ, M_δ and λ_δ such that, for any $\mathbf{z}, \mathbf{z}'$ in $\mathcal{D}$, and for all $t \geq 0$*

$$\|\langle \mathbf{Z} \rangle(\mathbf{z}, \varepsilon)\| \leq \overline{M} \, , \tag{11a}$$

$$\|\mathbf{Z}(t, \mathbf{z}, \varepsilon) - \mathbf{Z}(t, \mathbf{z}', \varepsilon)\| \leq \lambda \|\mathbf{z} - \mathbf{z}'\| \, , \tag{11b}$$

$$\|\Delta\mathbf{Z}(t, \mathbf{z}, \varepsilon)\| \leq M_\delta \, , \tag{11c}$$

$$\|\Delta\mathbf{Z}(t, \mathbf{z}, \varepsilon) - \Delta\mathbf{Z}(t, \mathbf{z}', \varepsilon)\| \leq \lambda_\delta \|\mathbf{z} - \mathbf{z}'\| \, , \tag{11d}$$

where $\Delta\mathbf{Z} = \mathbf{Z} - \langle \mathbf{Z} \rangle$, then the statement (8) *is modified as follows:*

$$\|\mathbf{z}(t) - \mathbf{u}(t)\| \leq \varepsilon \, \mathrm{e}^{\lambda L} \, T \left(\frac{L}{2} \lambda_\delta \overline{M} + M_\delta \right) \, . \tag{12}$$

We notice that Proposition 1 always provides better estimates than Theorem 1, since

$$\varepsilon \, \mathrm{e}^{\lambda L} \, T \left(\frac{L}{2} \lambda_\delta \overline{M} + M_\delta \right) \leq \varepsilon \, \mathrm{e}^{\lambda L} \, T \left(L\lambda M + 2M \right) .$$

The proof of this proposition is identical to the proof of Theorem 1. We notice that the constant $\overline{M}$ is the maximum value of the averaged flow, while M is the maximum value of the flow itself in Theorem 1.

2.2. Application to a one-dimensional soft-Coulomb potential

We consider a soft-Coulomb potential in one dimension. Hamiltonian (3) becomes

$$H(x,p,t) = \frac{p^2}{2} + V\left(x + \frac{E(t)}{\omega^2}\right) ,$$

where

$$V\left(x + \frac{E(t)}{\omega^2}\right) = -\frac{1}{\sqrt{(x + q\cos(\omega t))^2 + a^2}} ,$$

with $q = E_0/\omega^2$ the quiver radius.

Rescaling procedure : We cast the equations of motion associated with Hamiltonian H in the form (5) in order to apply Theorem 1. We define $\phi = \omega t$ as the rescaled time/evolution parameter, and rescale position and momentum as

$$x = C_x\,\chi\,, \qquad p = C_p\,\xi\,.$$

Inserting these rescalings, the equations of motion become

$$\frac{\mathrm{d}\chi}{\mathrm{d}\phi} = \frac{C_p}{\omega\,C_x}\,\xi\,, \qquad \frac{\mathrm{d}\xi}{\mathrm{d}\phi} = -\frac{1}{\omega\,C_p C_x^2}\,W'(\chi + (q/C_x)\cos\phi)\,,$$

where W is the rescaled potential:

$$V(x + q\cos\omega t) = \frac{1}{C_x}\,W\left(\chi + (q/C_x)\cos\phi\right)\,.$$

So, the factors C_x, C_p and the small parameter ε of Bogolyubov's theorem must satisfy the relation

$$\frac{C_p}{\omega\,C_x} = \frac{1}{\omega\,C_p C_x^2} = \varepsilon\,.$$

The choice of C_x is guided by the fact that the rescaling of time and space is ruled by the external field, which governs the fast dynamics. That is why we choose C_x to be the quiver radius q. We then have

$$\varepsilon = \frac{1}{\omega q^{3/2}} = \frac{\omega^2}{E_0^{3/2}}\,, \qquad \phi = \omega t\,, \qquad \chi = \frac{x}{q}\,, \qquad \xi = \sqrt{q}\,p\,, \tag{13}$$

where χ and ξ are the rescaled position and momentum, respectively. The equations of motion in the rescaled variables are:

$$\frac{\mathrm{d}\chi}{\mathrm{d}\phi} = \varepsilon\,\xi\,, \qquad \frac{\mathrm{d}\xi}{\mathrm{d}\phi} = -\varepsilon\,W'(\chi + \cos\phi)\,,$$

where $\mathbf{z} = (\chi, \xi)$ in the notations of Theorem 1 and

$$W(\chi) = -\frac{1}{\sqrt{\chi^2 + \alpha^2}} \,,$$

where $\alpha = a/q$ is the rescaled softening parameter.

Using the KH approximation, the rescaled KH potential is denoted W_{KH} and reads

$$W_{\mathrm{KH}}(\chi) = -\frac{1}{2\pi} \int_0^{2\pi} \frac{\mathrm{d}\phi}{\sqrt{(\chi + \cos\phi)^2 + \alpha^2}} = -\left\langle \frac{1}{\sqrt{(\chi + \cos\phi)^2 + \alpha^2}} \right\rangle .$$

Therefore the rescaled potential is related to the unrescaled potential by $W_{\mathrm{KH}}(\chi) = qV_{\mathrm{KH}}(q\chi)$. The averaged trajectory $\mathbf{u} = (\overline{\chi}, \overline{\xi})$ obeys the equations:

$$\frac{\mathrm{d}\overline{\chi}}{\mathrm{d}\phi} = \varepsilon\, \overline{\xi} \,, \qquad \frac{\mathrm{d}\overline{\xi}}{\mathrm{d}\phi} = -\varepsilon\, W'_{\mathrm{KH}}(\overline{\chi}).$$

The flows involved in Theorem 1 are given by

$$\begin{aligned} \mathbf{Z}(\phi, \mathbf{z}) &= (\xi \,, -W'\,(\chi + \cos\phi)) \ , \\ \langle \mathbf{Z} \rangle(\mathbf{u}) &= \left(\overline{\xi} \,, -W'_{\mathrm{KH}}(\overline{\chi})\right) . \end{aligned}$$

In what follows, we use the 1-norm:

$$\|\mathbf{z}\| = |\chi| + |\xi|.$$

Other choices would give similar results, thanks to the equivalence of norms in finite dimension.

The constants M, λ, $\overline{M}$, M_δ and λ_δ appearing in Theorem 1 and Proposition 1 are expressed as functions of the bounds for the absolute values of ξ, W, W_{KH} and their derivatives. We denote these by

$$|\xi| \le B_p \,, \qquad |W(\chi + \cos\phi)| \le B_W \,, \qquad \left|W^{(n)}(\chi + \cos\phi)\right| \le B_{W,n} \,,$$
$$|W_{\mathrm{KH}}(\chi)| \le B_{\langle W \rangle} \,, \qquad \left|W^{(n)}_{\mathrm{KH}}(\chi)\right| \le B_{\langle W \rangle, n} \,.$$

These bounds have to hold for any $\phi \ge 0$ and will depend on the region (χ, ξ) of phase space where the trajectory evolves. We have then

$$M = \overline{M} = B_p + B_{W,1} \,, \quad \lambda = \max\left(1, B_{W,2}\right) \,, \quad M_\delta = B_{W,1} \,, \quad \lambda_\delta = B_{W,2} \,.$$

From Theorem 1, for all $0 \leq \phi \leq L/\varepsilon$, we have:

$$|\chi(\phi) - \overline{\chi}(\phi)| + |\xi(\phi) - \overline{\xi}(\phi)| \leq \varepsilon\, \mathrm{e}^{\lambda L}\, 2\pi\, (L\lambda M + 2M) \ ,$$

and, from Proposition 1,

$$|\chi(\phi) - \overline{\chi}(\phi)| + |\xi(\phi) - \overline{\xi}(\phi)| \leq \varepsilon\, \mathrm{e}^{\lambda L}\, 2\pi \left(\frac{L}{2} \lambda_\delta \overline{M} + M_\delta \right) \ .$$

To deduce an estimate for the original variables x_e, p_e and t, Eqs. (2) and (13) are used, i.e.,

$$x_\mathrm{e}(t) = q[\chi(\omega t) + \cos(\omega t)] \ , \qquad p_\mathrm{e}(t) = \frac{1}{\sqrt{q}} \left[\xi(\omega t) - \frac{1}{\varepsilon} \sin(\omega t) \right] \ ,$$

together with the analogous relations for the reconstructed trajectories of the averaged system (ruled by the time-independent Hamiltonian (4)), which are denoted by $\overline{x}_\mathrm{e}(t), \overline{p}_\mathrm{e}(t)$:

$$\overline{x}_\mathrm{e}(t) = q[\overline{\chi}(\omega t) + \cos(\omega t)] \ , \qquad \overline{p}_\mathrm{e}(t) = \frac{1}{\sqrt{q}} \left[\overline{\xi}(\omega t) - \frac{1}{\varepsilon} \sin(\omega t) \right] \ .$$

We notice that, given these expressions, we have $|x_\mathrm{e}(t) - \overline{x}_\mathrm{e}(t)| = |x(t) - \overline{x}(t)|$ and $|p_\mathrm{e}(t) - \overline{p}_\mathrm{e}(t)| = |p(t) - \overline{p}(t)|$.

We distinguish the case where χ takes values on the whole real line, and the case where χ stays larger than a prescribed value. In each case, we estimate the bounds Bs, using energy conservation on averaged trajectories

$$E_\mathrm{KH} = \frac{1}{q} \left[\frac{\xi(0)^2}{2} + W_\mathrm{KH}(\chi(0)) \right] = \frac{1}{q} \left[\frac{\overline{\xi}(\phi)^2}{2} + W_\mathrm{KH}(\overline{\chi}(\phi)) \right]$$

for all ϕ. We notice that the bounds M, $\overline{M}$, M_δ and λ_δ are obtained from the averaged trajectory $\mathbf{u}(t)$, which allows the use of the energy conservation for their estimates (see, e.g., Eqs. (9)-(10)).

2.2.1. Case $\chi \in \mathbb{R}$

If we consider the whole domain $\mathbb{R}^2$ for (χ, ξ), then

$$B_W = |W(0)| = \frac{1}{\alpha} \ , \quad B_{W,1} = \left| W'\left(\frac{\alpha}{\sqrt{2}} \right) \right| = \frac{2}{3\sqrt{3}\,\alpha^2} \ , \quad B_{W,2} = |W''(0)| = \frac{1}{\alpha^3} \ , \tag{14}$$

and since we do not have an explicit expression for $W_{\rm KH}$ we take $B_{\langle W\rangle,n} = B_{W,n}$. For bound states ($E_{\rm KH} < 0$), energy conservation on averaged trajectories implies that

$$|\overline{\xi}| \leq \sqrt{2\,(qE_{\rm KH} + B_W)} \leq \sqrt{2B_W} = B_p\,.$$

In sum: Theorem 1 gives, for all $0 \leq t < t_{\max}$:

$$\frac{1}{q}\,|x_{\rm e}(t) - \overline{x}_{\rm e}(t)| + \sqrt{q}\,|p_{\rm e}(t) - \overline{p}_{\rm e}(t)| \leq \frac{T}{q^{3/2}}\,{\rm e}^{\lambda\, t_{\max}\, q^{-3/2}} \left(\frac{t_{\max}\lambda M}{q^{3/2}} + 2M\right)\,,$$

where

$$M = \frac{2}{3\sqrt{3}\,\alpha^2} + \sqrt{\frac{2}{\alpha}}\,, \qquad \lambda = \max\left(1, \frac{1}{\alpha^3}\right)\,.$$

If Proposition 1 is used, we have

$$\frac{1}{q}\,|x_{\rm e}(t) - \overline{x}_{\rm e}(t)| + \sqrt{q}\,|p_{\rm e}(t) - \overline{p}_{\rm e}(t)| \leq \frac{T}{q^{3/2}}\,{\rm e}^{\lambda\, t_{\max}\, q^{-3/2}} \left(\frac{t_{\max}\lambda_\delta M}{2\,q^{3/2}} + M_\delta\right)\,,$$

with

$$M_\delta = \frac{2}{3\sqrt{3}\,\alpha^2}\,, \qquad \lambda_\delta = \frac{1}{\alpha^3}\,.$$

For typical values of laser intensity $I = 5 \times 10^{15}$ W cm^{-2}, frequency $\omega = 0.0584$, $a = 1$ and $t_{\max} = 5$ laser cycles, we have $M \simeq M_\delta \simeq 4.7 \times 10^3$, $\lambda = \lambda_\delta \simeq 1.4 \times 10^6$. So, the preceding bounds are mostly useless when evaluated at realistic values of the parameters. This is due to the fact that the electron is allowed to go arbitrarily close to the ionic core: the value of the bounds (14) is reached when the position $q(\chi + \cos\phi)$ of the electron is zero or close to zero.

Remark: The same bounds apply to the case where the trajectory of the averaged system lies near the local minimum χ_* of the KH potential, i.e., to the case $1 - \mu \leq |\chi| \leq 1$, with $\mu \in [0, 1]$. This is due to the fact that the absolute value of the argument $\chi + \cos\phi$ of the potential W takes all the values from 0 to $2 - \mu$ as ϕ varies, so that the absolute maxima of $|W|$ and $|W''|$ given by Eq. (14) are attained. In particular, the exponential behavior of the estimate is mostly identical for $\chi \in \mathbb{R}$ and for $|\chi| \in [1 - \mu, 1]$.

2.2.2. Case $|\chi| \geq 1+\mu$

We now consider the domain $\mathcal{D}_\mu$ defined by $|\chi| \geq 1+\mu$ (and arbitrary ξ) with μ a fixed positive number. The bounds are now:

$$B_W = |W(\mu)| = \frac{1}{\sqrt{\mu^2+\alpha^2}} \quad \text{for any} \quad \mu > 0\,,$$
$$B_{W,1} = |W'(\mu)| = \frac{\mu}{(\mu^2+\alpha^2)^{3/2}} \quad \text{if} \quad \mu \geq \frac{\alpha}{\sqrt{2}}\,,$$
$$B_{W,2} = |W''(\mu)| = \frac{2\mu^2-\alpha^2}{(\mu^2+\alpha^2)^{5/2}} \quad \text{if} \quad \mu \geq \sqrt{\frac{3}{2}}\,\alpha\,.$$

In sum: For x_e and $\overline{x}_\mathrm{e}$ satisfying the condition $|x_\mathrm{e}(t)|, |\overline{x}_\mathrm{e}(t)| \geq \mu q$ and $\mu \geq \alpha\sqrt{3/2}$, Theorem 1 gives, for all $0 \leq t < t_\mathrm{max}$:

$$\frac{1}{q}\,|x_\mathrm{e}(t)-\overline{x}_\mathrm{e}(t)| + \sqrt{q}\,|p_\mathrm{e}(t)-\overline{p}_\mathrm{e}(t)| \leq \frac{T}{q^{3/2}}\,\mathrm{e}^{\lambda\, t_\mathrm{max}\, q^{-3/2}}\left(\frac{t_\mathrm{max}\lambda M}{q^{3/2}} + 2M\right), \tag{15}$$

where

$$M = \frac{\mu}{(\mu^2+\alpha^2)^{3/2}} + \frac{\sqrt{2}}{(\mu^2+\alpha^2)^{1/4}}\,, \qquad \lambda = \max\left(1, \frac{2\mu^2-\alpha^2}{(\mu^2+\alpha^2)^{5/2}}\right).$$

If Proposition 1 is used, we have

$$\frac{1}{q}\,|x_\mathrm{e}(t)-\overline{x}_\mathrm{e}(t)| + \sqrt{q}\,|p_\mathrm{e}(t)-\overline{p}_\mathrm{e}(t)| \leq \frac{T}{q^{3/2}}\,\mathrm{e}^{\lambda\, t_\mathrm{max}\, q^{-3/2}}\left(\frac{t_\mathrm{max}\lambda_\delta M}{2\,q^{3/2}} + M_\delta\right), \tag{16}$$

with

$$M_\delta = \frac{\mu}{(\mu^2+\alpha^2)^{3/2}}\,, \qquad \lambda_\delta = \frac{2\mu^2-\alpha^2}{(\mu^2+\alpha^2)^{5/2}}\,.$$

In Fig. 1, we represent the right-hand side of the estimates (15)-(16) as a function of μ (respectively, blue and magenta curves). We notice that the estimate given by Eq. (16) is much more accurate for large values of μ, i.e., when the trajectory is far away from the core (of the order of several quiver radii). Close to the ionic core, both estimates are equivalently inaccurate.

2.3. Discussion

Is the Kramers-Henneberger reduction valid? In order to answer this question, Bogolyubov's averaging theorem and a variant of it were used.

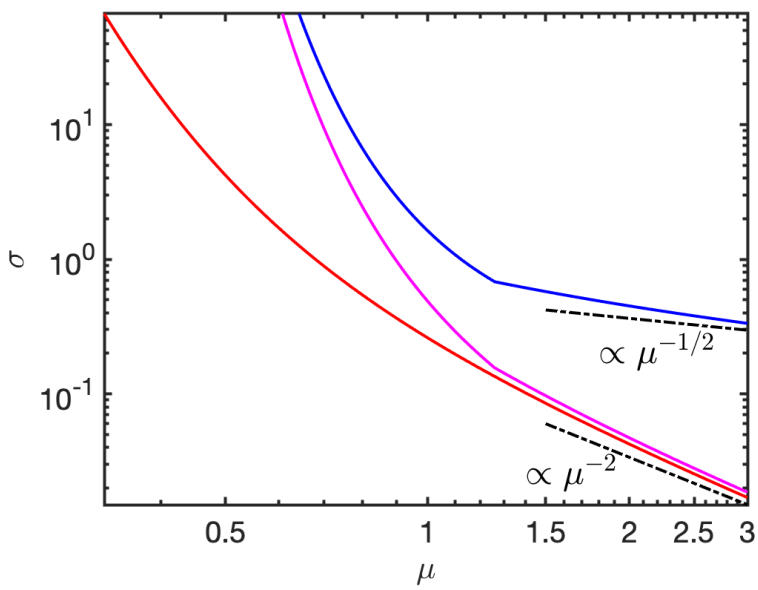

Figure 1: Loglog plot of $\sigma = |x_{\rm e}(t) - \overline{x}_{\rm e}(t)|/q + \sqrt{q}\,|p_{\rm e}(t) - \overline{p}_{\rm e}(t)|$ as a function of μ at $I = 5 \times 10^{15}$ W cm^{-2}, $\omega = 0.0584$, $a = 1$ and $t_{\rm max} = 5$ laser cycles. Blue line: from Theorem 1, Eq. (15). Magenta line: from Proposition 1, Eq. (16). Red line: from Theorem 2, Eq. (23).

In the estimates presented in this section, the most damaging term is the exponential term, $\mathrm{e}^{\lambda t_{\max} q^{-3/2}}$ which can be rewritten as $\mathrm{e}^{\varepsilon\lambda\omega t_{\max}}$. This term is present in both Theorem 1 and Proposition 1 with the same exponent λ. This exponential term becomes $\mathrm{e}^{2\pi\varepsilon\lambda n}$, where n is the duration of the laser pulse in laser cycles. This term remains of order one if $\varepsilon\lambda$ is much smaller than 1. The best-case scenario is when the electron is relatively far away from the ionic core, where λ is of order 1 (notice that it is even larger close to the ionic core). This leads to an upper bound for ε which translates into a lower bound for E_0:

$$E_0 \gg \omega^{4/3}(2\pi n)^{2/3}.$$

For an infrared laser field with $\omega = 0.0584$ with a duration of five laser cycles, this means an intensity greater than 1.8×10^{15} W cm^{-2}, which is much larger than what is typically considered for KH simulations. Following Ref. [20], an alternative way to make this exponential term small is to reduce the size of the pulse, namely n, and considering

$$n \lesssim \frac{1}{2\pi\lambda\sqrt{\varepsilon}}\,.$$

For $\omega = 0.0584$ and an intensity of 5×10^{15} W cm^{-2}, this gives n less than approximately one laser cycle, which is not realistic (and it is even worse for smaller intensities). Therefore the only way to reduce the effects of the exponential term is to have parameters of the field so that ε is typically smaller than 1/10, and being relatively far away from the ionic core, typically more than a couple of quiver radii.

Does this mean that the Kramers-Henneberger reduction is invalid? Of course, not. This just means that estimates following Bogolyubov's theorem are not suited to address this question, at least at a distance closer than one quiver radius of the ionic core, since they neglect the true dynamics in phase space, chaotic in nature close to the bounded region and rather regular in large portions of phase space away from it. More importantly, the estimates are very general, and not tailored to Hamiltonian systems. To illustrate this point, we consider the following simple example for $\mathbf{z} = (x, p)$ and Hamiltonian $H(t, \mathbf{z}, \varepsilon) = \varepsilon(p^2/2 - x\cos t)$:

$$\mathbf{Z}(t, \mathbf{z}, \varepsilon) = \varepsilon(p\ , \cos t)\,, \tag{17a}$$

$$\langle \mathbf{Z} \rangle(\mathbf{z}, \varepsilon) = \varepsilon(p\ , 0)\,. \tag{17b}$$

The constant λ in Theorem 1 and Proposition 1 is equal to 1, so that

$$\|\mathbf{z}(t) - \mathbf{u}(t)\| \leq 2\pi\varepsilon M(\varepsilon t_{\max} + 2)\mathrm{e}^{\varepsilon t_{\max}} ,$$

for all $|p| \leq M - 1$ and $t \leq t_{\max}$. However, in this case, the Hamiltonian flow (17) is easily integrated, and the estimate computed exactly:

$$\|\mathbf{z}(t) - \mathbf{u}(t)\| = \varepsilon|\sin t| + \varepsilon^2|1 - \cos t| \leq \varepsilon(1 + 2\varepsilon) ,$$

for all times. Instead of an exponential growth of the distance, this distance remains bounded in time. Therefore, we see on one of the simplest examples that estimates following Bogolyubov's theorem are far from being optimal. This is expected since these estimates are obtained with very general assumptions on the flow regardless of, e.g., its geometrical structure.

Below, we modify Bogolyubov's theorem so that the estimates are better adapted to Hamiltonian systems.

3. Modification of Bogolyubov's averaging theorem

3.1. Statement of the theorem

First, we separate the variables into positions and momenta:

$$\mathbf{z} = (\mathbf{r} , \mathbf{p}) .$$

The equations of motion are of the form

$$\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}t} = \varepsilon \, (\nabla_{\mathbf{p}} H , -\nabla_{\mathbf{r}} H) = \varepsilon \, \mathbf{Z}(t, \mathbf{z}, \varepsilon) .$$

The trajectory $\mathbf{z}(t)$ is compared to the trajectory $\mathbf{u}(t) = (\overline{\mathbf{r}}(t) , \overline{\mathbf{p}}(t))$ of the averaged system

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = \varepsilon \, (\nabla_{\mathbf{p}} \langle H \rangle , -\nabla_{\mathbf{r}} \langle H \rangle) = \varepsilon \, \langle \mathbf{Z} \rangle (\mathbf{u}, \varepsilon) .$$

Following Bogolyubov's theorem, we consider that the two trajectories have the same initial condition $\mathbf{z}(0) = \mathbf{u}(0) = \mathbf{z}_0$. We estimate the differences in the positions and in momenta, i.e., the vector

$$\begin{pmatrix} \|\mathbf{r}(t) - \overline{\mathbf{r}}(t)\| \\ \|\mathbf{p}(t) - \overline{\mathbf{p}}(t)\| \end{pmatrix} = \sum_{i=1}^{d} \begin{pmatrix} |r_i(t) - \overline{r}_i(t)| \\ |p_i(t) - \overline{p}_i(t)| \end{pmatrix} ,$$

where d is the dimension of configuration space.

Two Cauchy problems are compared:

$$\mathrm{I}\ :\ \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \varepsilon\,\nabla_{\mathbf{p}} H(t,\mathbf{r},\mathbf{p},\varepsilon)\ ,\ \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = -\varepsilon\,\nabla_{\mathbf{r}} H(t,\mathbf{r},\mathbf{p},\varepsilon)\,,$$
$$\mathrm{II}\ :\ \frac{\mathrm{d}\overline{\mathbf{r}}}{\mathrm{d}t} = \varepsilon\,\nabla_{\mathbf{p}} \langle H\rangle(\overline{\mathbf{r}},\overline{\mathbf{p}},\varepsilon)\ ,\ \frac{\mathrm{d}\overline{\mathbf{p}}}{\mathrm{d}t} = -\varepsilon\,\nabla_{\mathbf{r}} \langle H\rangle(\overline{\mathbf{r}},\overline{\mathbf{p}},\varepsilon)\,,$$

with $\mathbf{r}(0) = \overline{\mathbf{r}}(0) = \mathbf{r}_0$ and $\mathbf{p}(0) = \overline{\mathbf{p}}(0) = \mathbf{p}_0$. Here $H(t,\mathbf{r},\mathbf{p},\varepsilon)$ is defined in a domain $\mathcal{D}$ of $(\mathbf{r},\mathbf{p})$, and is a periodic function of t, of period T.

We suppose that, for $t \geq 0$ and $0 \leq \varepsilon < \varepsilon_1$, it is possible to define two-dimensional vectors

$$\overline{\mathbf{M}} = \begin{pmatrix} M_p \\ M_r \end{pmatrix}\ , \qquad \mathbf{M}_\delta = \begin{pmatrix} M_p^\delta \\ M_r^\delta \end{pmatrix}\ , \tag{18}$$

and 2×2 matrices

$$\mathbb{\Lambda} = \begin{pmatrix} \lambda_{rp} & \lambda_{pp} \\ \lambda_{rr} & \lambda_{rp} \end{pmatrix}\ , \qquad \mathbb{\Lambda}_\delta = \begin{pmatrix} \lambda_{rp}^\delta & \lambda_{pp}^\delta \\ \lambda_{rr}^\delta & \lambda_{rp}^\delta \end{pmatrix}\ , \tag{19}$$

such that the following conditions hold for $t \geq 0$, for any $\mathbf{z} = (\mathbf{r}\,,\mathbf{p})$ in $\mathcal{D}$:

$$\max_{1\leq i\leq d} |\nabla_{r_i}\langle H\rangle| \leq \frac{M_r}{d}\,, \qquad \max_{1\leq i\leq d} |\nabla_{p_i}\langle H\rangle| \leq \frac{M_p}{d}\,, \tag{20a}$$

$$\max_{1\leq i\leq d} |\nabla_{r_i}\Delta H| \leq \frac{M_r^\delta}{d}\,, \qquad \max_{1\leq i\leq d} |\nabla_{p_i}\Delta H| \leq \frac{M_p^\delta}{d}\,, \tag{20b}$$

$$\max_{1\leq i,j\leq d} |\nabla_{r_i}\nabla_{r_j} H| \leq \frac{\lambda_{rr}}{d}\,,\ \max_{1\leq i,j\leq d} |\nabla_{p_i}\nabla_{p_j} H| \leq \frac{\lambda_{pp}}{d}\,,\ \max_{1\leq i,j\leq d} |\nabla_{p_i}\nabla_{r_j} H| \leq \frac{\lambda_{rp}}{d}\,, \tag{20c}$$

$$\max_{1\leq i,j\leq d} |\nabla_{r_i}\nabla_{r_j} \Delta H| \leq \frac{\lambda_{rr}^\delta}{d}\,,\ \max_{1\leq i,j\leq d} |\nabla_{p_i}\nabla_{p_j} \Delta H| \leq \frac{\lambda_{pp}^\delta}{d}\,,\ \max_{1\leq i,j\leq d} |\nabla_{p_i}\nabla_{r_j} \Delta H| \leq \frac{\lambda_{rp}^\delta}{d}\,, \tag{20d}$$

where

$$\Delta H(t,\mathbf{z},\varepsilon) = H(t,\mathbf{z},\varepsilon) - \langle H\rangle(\mathbf{z},\varepsilon)\,.$$

Remark: If we denote by $\mathbb{A}$, $\mathbb{A}_\delta$ the Hessian matrices of H, ΔH, and $\mathbb{J}_{2d}$ the symplectic matrix:

$$A_{ij} = \nabla_{z_i}\nabla_{z_j} H\,, \qquad A_{ij}^\delta = \nabla_{z_i}\nabla_{z_j}\Delta H\,, \qquad \mathbb{J}_{2d} = \begin{pmatrix} 0 & \mathbb{I}_d \\ -\mathbb{I}_d & 0 \end{pmatrix}\ ,$$

where $\mathbb{I}_d$ is the $d \times d$ identity matrix, we see that $\overline{\mathbf{M}}$, $\mathbf{M}_\delta$ give bounds for the absolute value of the vectors $\mathbb{J}_{2d}\nabla_{\mathbf{z}}\langle H\rangle$, $\mathbb{J}_{2d}\nabla_{\mathbf{z}}\Delta H$, while $\mathbb{\Lambda}$, $\mathbb{\Lambda}_\delta$ give bounds for the absolute value of $\mathbb{J}\mathbb{A}$, $\mathbb{J}\mathbb{A}_\delta$.

Theorem 2. *If the solution* $(\overline{\mathbf{r}}(t), \overline{\mathbf{p}}(t))$ *to the initial Cauchy problem for system II is defined for* $t \geq 0$ *and belongs to the domain* $\mathcal{D}$*, then for any* $L > 0$*, the following inequality holds*

$$\begin{pmatrix} \|\mathbf{r}(t) - \overline{\mathbf{r}}(t)\| \\ \|\mathbf{p}(t) - \overline{\mathbf{p}}(t)\| \end{pmatrix} \leq \mathrm{e}^{\mathbb{\Lambda}L}\varepsilon T \left[\frac{L}{2}\mathbb{\Lambda}_\delta\overline{\mathbf{M}} + \mathbf{M}_\delta\right], \tag{21}$$

for all $0 \leq t < L/\varepsilon$ *and* $0 \leq \varepsilon < \varepsilon_1$*, with* $\overline{\mathbf{M}}$*,* $\mathbf{M}_\delta$*,* $\mathbb{\Lambda}$ *and* $\mathbb{\Lambda}_\delta$ *defined by Eqs. (18), (19) and (20).*

Theorem 2 is essentially a matrix generalization of Theorem 1, or more precisely of Proposition 1, which improves the estimates as seen in Sec. 2.2.2. The constants $\overline{M}$, λ, M_δ and λ_δ of Eq. (11) are respectively replaced by two-dimensional vectors $\overline{\mathbf{M}}$, $\mathbf{M}_\delta$, and 2×2 matrices $\mathbb{\Lambda}$, $\mathbb{\Lambda}_\delta$.

The proof of Theorem 2 follows the same lines as that of Theorem 1. We write

$$\begin{aligned}
\begin{pmatrix} \|\mathbf{r}(t) - \overline{\mathbf{r}}(t)\| \\ \|\mathbf{p}(t) - \overline{\mathbf{p}}(t)\| \end{pmatrix} &= \varepsilon \begin{pmatrix} \| \int_0^t [\nabla_{\mathbf{p}} H(\tau, \mathbf{z}(\tau), \varepsilon) - \nabla_{\mathbf{p}}\langle H\rangle(\tau, \mathbf{u}(\tau), \varepsilon)]\mathrm{d}\tau\| \\ \| \int_0^t [\nabla_{\mathbf{r}} H(\tau, \mathbf{z}(\tau), \varepsilon) - \nabla_{\mathbf{r}}\langle H\rangle(\tau, \mathbf{u}(\tau), \varepsilon)]\mathrm{d}\tau\| \end{pmatrix} \\
\leq \varepsilon \begin{pmatrix} \int_0^t \|\nabla_{\mathbf{p}} H(\tau, \mathbf{z}(\tau), \varepsilon) - \nabla_{\mathbf{p}} H(\tau, \mathbf{u}(\tau), \varepsilon)\|\mathrm{d}\tau + \| \int_0^t \nabla_{\mathbf{p}}\Delta H(\tau, \mathbf{u}(\tau), \varepsilon)\mathrm{d}\tau\| \\ \int_0^t \|\nabla_{\mathbf{r}} H(\tau, \mathbf{z}(\tau), \varepsilon) - \nabla_{\mathbf{r}} H(\tau, \mathbf{u}(\tau), \varepsilon)\|\mathrm{d}\tau + \| \int_0^t \nabla_{\mathbf{r}}\Delta H(\tau, \mathbf{u}(\tau), \varepsilon)\mathrm{d}\tau\| \end{pmatrix} \\
\leq \varepsilon\mathbb{\Lambda} \begin{pmatrix} \int_0^t \|\mathbf{r}(\tau) - \overline{\mathbf{r}}(\tau)\|\mathrm{d}\tau \\ \int_0^t \|\mathbf{p}(\tau) - \overline{\mathbf{p}}(\tau)\|\mathrm{d}\tau \end{pmatrix} + \varepsilon\mathbf{C},
\end{aligned}$$

where $\mathbf{C}$ is defined by

$$\mathbf{C} = \sup_{t\in[0,L/\varepsilon]} \begin{pmatrix} \| \int_0^t \nabla_{\mathbf{p}}\Delta H(\tau, \mathbf{u}(\tau), \varepsilon)\mathrm{d}\tau\| \\ \| \int_0^t \nabla_{\mathbf{r}}\Delta H(\tau, \mathbf{u}(\tau), \varepsilon)\mathrm{d}\tau\| \end{pmatrix}.$$

We now use the following generalization of Grönwall's inequality:

Lemma 1 ([26]). *Let the vector* $\mathbf{a}(t)$ *and the nonnegative matrices* $\mathbb{G}(t)$ *and* $\mathbb{H}(t)$ *be functions of the scalar variable* t*. Assume that* $\mathbb{H}(t)\mathbb{G}(t)$ *and* $\int_{t_0}^t \mathbb{H}(s)\mathbb{G}(s)\mathrm{d}s$ *commute for* $t > t_0$*. If the inequality*

$$\boldsymbol{\sigma}(t) \leq \mathbf{a}(t) + \mathbb{G}(t)\int_{t_0}^t \mathbb{H}(\tau)\boldsymbol{\sigma}(\tau)\mathrm{d}\tau$$

holds for all $t > t_0$, then

$$\boldsymbol{\sigma}(t) \leq \mathbf{a}(t) + \mathbb{G}(t) \int_{t_0}^{t} \exp\left(\int_{\tau}^{t} \mathbb{H}(s)\mathbb{G}(s)\mathrm{d}s\right)\mathbb{H}(\tau)\mathbf{a}(\tau)\mathrm{d}\tau$$

holds for $t > t_0$.

In our case $\mathbf{a}(t) = \varepsilon\mathbf{C}$, $\mathbb{G}(t) = \varepsilon\mathbb{A}$ independent of t, and $\mathbb{H}(t)$ is the identity matrix. This leads to:

$$\begin{pmatrix} \|\mathbf{r}(t) - \overline{\mathbf{r}}(t)\| \\ \|\mathbf{p}(t) - \overline{\mathbf{p}}(t)\| \end{pmatrix} \leq \left[\mathbb{I}_2 + \varepsilon\mathbb{A}\int_0^t \mathrm{d}\tau\, \mathrm{e}^{(t-\tau)\varepsilon\mathbb{A}}\right]\varepsilon\mathbf{C} = \mathrm{e}^{\varepsilon t\mathbb{A}}\varepsilon\mathbf{C}\,.$$

Here we clearly notice that the estimates for $\mathbf{C}$ are obtained from the trajectories $\mathbf{u}(t)$ in the averaged system.

We evaluate the vector $\mathbf{C}$ in the case where $H(t, \mathbf{z}, \varepsilon)$ is periodic in t, i.e., $H(t, \mathbf{z}, \varepsilon) = H(t + T, \mathbf{z}, \varepsilon)$ for all t. Proceeding in the same way as in the proof of Theorem 1, we partition the interval $[0, t]$: For $nT \leq t < (n + 1)T$, i.e., $n = \lfloor t/T \rfloor$, we get (say, for the first component):

$$\left\|\int_0^t \nabla_{\mathbf{p}}\Delta H(\tau, \mathbf{u}(\tau), \varepsilon)\mathrm{d}\tau\right\| \leq \sum_{i=1}^{n}\int_{(i-1)T}^{iT} \|\nabla_{\mathbf{p}}\Delta H(\tau, \mathbf{u}(\tau), \varepsilon) - \nabla_{\mathbf{p}}\Delta H(\tau, \mathbf{u}_i, \varepsilon)\|\mathrm{d}\tau$$
$$+ \int_{nT}^{t} \|\nabla_{\mathbf{p}}\Delta H(\tau, \mathbf{u}(\tau), \varepsilon)\|\,\mathrm{d}\tau\,,$$

where $\mathbf{u}_i = \mathbf{u}((i - 1)T)\,,\ i = 1, \ldots, n$ denote the values of $\mathbf{u}$ at each period. It follows that

$$\left\|\int_0^t \nabla_{\mathbf{p}}\Delta H(\tau, \mathbf{u}(\tau), \varepsilon)\mathrm{d}\tau\right\| \leq \sum_{i=1}^{n}\int_{(i-1)T}^{iT} [\lambda_{rp}^{\delta}\|\overline{\mathbf{r}}(\tau) - \overline{\mathbf{r}}_i\| + \lambda_{pp}^{\delta}\|\overline{\mathbf{p}}(\tau) - \overline{\mathbf{p}}_i\|]\mathrm{d}\tau + M_p^{\delta}T\,.$$

Using the fact that for $\tau \in [(i - 1)T, iT]$

$$\mathbf{u}(\tau) = \mathbf{u}_i + \int_{(i-1)T}^{\tau} \varepsilon\,\langle\mathbf{Z}\rangle(\mathbf{u}(s), \varepsilon)\,\mathrm{d}s\,,$$

we get

$$\|\overline{\mathbf{r}}(\tau) - \overline{\mathbf{r}}_i\| \leq \varepsilon M_p(\tau - (i - 1)T)\,, \qquad \|\overline{\mathbf{p}}(\tau) - \overline{\mathbf{p}}_i\| \leq \varepsilon M_r(\tau - (i - 1)T)\,,$$

so that

$$\left\| \int_0^t \nabla_{\mathbf{p}} \Delta H(\tau, \mathbf{u}(\tau), \varepsilon) \mathrm{d}\tau \right\| \leq \varepsilon n \frac{T^2}{2} (\lambda_{rp}^\delta M_p + \lambda_{pp}^\delta M_r) + M_p^\delta T,$$
$$\leq T \left[\frac{\varepsilon t}{2} (\lambda_{rp}^\delta M_p + \lambda_{pp}^\delta M_r) + M_p^\delta \right],$$

for all $t \leq L/\varepsilon$. Performing the same estimate for the second component leads to

$$\mathbf{C} = T \left[\frac{L}{2} \mathbb{\Lambda}_\delta \overline{\mathbf{M}} + \mathbf{M}_\delta \right],$$

from which Eq. (21) follows.

Remark : Thanks to the fact that the elements on the diagonal of $\mathbb{\Lambda}$ are identical (which comes from the Hamiltonian character of the dynamics) the exponential $\mathrm{e}^{\mathbb{\Lambda} L}$ has a closed expression. For example, in the usual case where Hamiltonian H is at most quadratic in momenta (i.e., $\lambda_{pp} = 1$) we have:

$$\mathrm{e}^{\mathbb{\Lambda} L} = \mathrm{e}^{\lambda_{rp} L} \begin{pmatrix} \cosh(\sqrt{\lambda_{rr}} L) & \frac{1}{\sqrt{\lambda_{rr}}} \sinh(\sqrt{\lambda_{rr}} L) \\ \sqrt{\lambda_{rr}} \sinh(\sqrt{\lambda_{rr}} L) & \cosh(\sqrt{\lambda_{rr}} L) \end{pmatrix}. \tag{22}$$

Equation (21) is the matrix equivalent of the scalar expression (12) of Proposition 1. The advantage of the matrix formulation is evident in the treatment of the simple example (17) considered in section 2.3. We have in this case:

$$\mathbb{\Lambda} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \overline{\mathbf{M}} = \begin{pmatrix} M_p \\ 0 \end{pmatrix}, \qquad \mathbb{\Lambda}_\delta = \mathbb{0}, \qquad \mathbf{M}_\delta = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

In this example, we notice that the value of M_p is irrelevant since $\overline{\mathbf{M}}$ is multiplied by the null matrix. Theorem 2 gives, for all $t \leq t_{\max}$:

$$|x(t) - \overline{x}(t)| \leq 2\pi\varepsilon^2 t_{\max}, \quad |p(t) - \overline{p}(t)| \leq 2\pi\varepsilon,$$

to be compared with the estimate coming from the exact solution:

$$|x(t) - \overline{x}(t)| = \varepsilon^2 |1 - \cos t| \leq 2\varepsilon^2, \quad |p(t) - \overline{p}(t)| = \varepsilon |\sin t| \leq \varepsilon.$$

As it has been pointed out in section 2.3, Theorem 1 gives:

$$|x(t) - \overline{x}(t)| + |p(t) - \overline{p}(t)| \leq 2\pi\varepsilon M(\varepsilon t_{\max} + 2)\mathrm{e}^{\varepsilon t_{\max}},$$

and the modified version of Proposition 1 (with $\lambda_\delta = 0$, $M_\delta = 1$) leads to :

$$|x(t) - \overline{x}(t)| + |p(t) - \overline{p}(t)| \leq 2\pi\varepsilon M \mathrm{e}^{\varepsilon t_{\max}} ,$$

for all $|p| \leq M - 1$ and $t \leq t_{\max}$. We notice that the matrix formulation in Theorem 2 allowed the removal of the exponentially divergent term for this specific case, since $\mathrm{e}^{t\mathbb{A}}$ is not necessarily exponential in time when $\mathbb{A}$ is a matrix. Next, we consider the running example of a one-dimensional soft-Coulomb potential to figure out if Theorem 2 improves the estimates and allows us to conclude on the validity of the Kramers-Henneberger approximation.

3.2. Application to a one-dimensional soft-Coulomb potential

For a one-dimensional soft-Coulomb potential, the matrices and vectors appearing in Theorem 2 are given by:

$$\mathbb{A} = \begin{pmatrix} 0 & 1 \\ B_{W,2} & 0 \end{pmatrix} , \quad \overline{\mathbf{M}} = \begin{pmatrix} B_p \\ B_{W,1} \end{pmatrix} , \quad \mathbb{A}_\delta = \begin{pmatrix} 0 & 0 \\ B_{W,2} & 0 \end{pmatrix} , \quad \mathbf{M}_\delta = \begin{pmatrix} 0 \\ B_{W,1} \end{pmatrix} ,$$

which gives, using Eqs. (21)-(22):

$$\begin{pmatrix} |\chi(\phi) - \overline{\chi}(\phi)| \\ |\xi(\phi) - \overline{\xi}(\phi)| \end{pmatrix} \leq 2\pi\varepsilon \left(\frac{L}{2} B_p B_{W,2} + B_{W,1} \right) \begin{pmatrix} \frac{1}{\sqrt{B_{W,2}}} \sinh(\sqrt{B_{W,2}} L) \\ \cosh(\sqrt{B_{W,2}} L) \end{pmatrix} . \tag{23}$$

Going back to the original variables x_e, p_e and t, and taking E_0/ω as the rescaling factor for momenta, we have

$$\frac{1}{q} \, |x_\mathrm{e}(t) - \overline{x}_\mathrm{e}(t)| \leq \frac{T}{q^{3/2}} \left(\frac{t_{\max}}{2q^{3/2}} B_p B_{W,2} + B_{W,1} \right) \frac{1}{\sqrt{B_{W,2}}} \sinh \left(\frac{t_{\max}\sqrt{B_{W,2}}}{q^{3/2}} \right) , \tag{24}$$

$$\frac{\omega}{E_0} \, |p_\mathrm{e}(t) - \overline{p}_\mathrm{e}(t)| \leq \frac{\varepsilon T}{q^{3/2}} \left(\frac{t_{\max}}{2q^{3/2}} B_p B_{W,2} + B_{W,1} \right) \cosh \left(\frac{t_{\max}\sqrt{B_{W,2}}}{q^{3/2}} \right) . \tag{25}$$

We first notice that there is still an exponential behavior in time for $t_{\max}$ large, but the exponent is $\sqrt{B_{W,2}}$ compared with the exponent for Theorem 1 and Proposition 1, $\max(1, B_{W,2})$. This reduction of exponent is crucial for small μs, and is one of the main ingredients for the improvement of the estimates. In Fig. 1, we have plotted the bounds σ for $|x_\mathrm{e}(t) - \overline{x}_\mathrm{e}(t)| / q +$

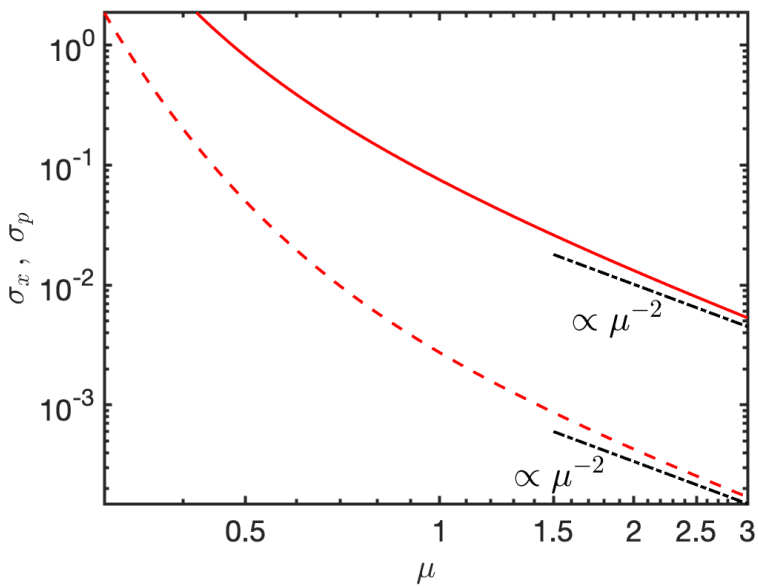

Figure 2: Loglog plot of $\sigma_x = \frac{\omega^2}{E_0}\ |x_{\rm e}(t) - \overline{x}_{\rm e}(t)|$ (red solid line) and $\sigma_p = \frac{\omega}{E_0}\ |p_{\rm e}(t) - \overline{p}_{\rm e}(t)|$ (red dash-dotted line) given by equations Eqs. (24)-(25) as a function of μ for $I = 5 \times 10^{15}$ W cm^{-2}, $\omega = 0.0584$, $a = 1$ and $t_{\rm max} = 5$ laser cycles.

$\sqrt{q}\,|p_{\rm e}(t) - \overline{p}_{\rm e}(t)|$ where the two contributions are obtained using Eqs. (24)-(25). We notice that for large μ the estimates obtained from Theorem 2 appear to be equivalent to the ones of Proposition 1. The main difference manifests itself for small values of μ, where the estimates improve closer to the ionic core.

In order to be more quantitative and compare the results of Theorem 1, Proposition 1 and Theorem 2 for the one-dimensional soft-Coulomb potential, we compute the estimates for large and small distances of the electron trajectory from the ionic core. This allows us to better understand the different behaviours displayed on Fig. 1 and identify quantitatively the improvement of each estimate. For each case, we give the leading term in the expression

$$\sigma = \max_{t \leq t_{\rm max}} \left(\frac{1}{q}\ |x_{\rm e}(t) - \overline{x}_{\rm e}(t)| + \sqrt{q}\ |p_{\rm e}(t) - \overline{p}_{\rm e}(t)| \right) .$$

Large distance: $\mu \gg 1$.

$$\sigma^{\mathrm{Th1}} \simeq \frac{T}{q^{3/2}}\, \mathrm{e}^{t_{\max}\, q^{-3/2}} \left(\frac{t_{\max}}{q^{3/2}} + 2 \right) \sqrt{\frac{2}{\mu}}\,,$$

$$\sigma^{\mathrm{Prop1}} \simeq \frac{T}{q^{3/2}}\, \mathrm{e}^{t_{\max}\, q^{-3/2}} \frac{1}{\mu^2}\,, \tag{26}$$

$$\sigma^{\mathrm{Th2}} \simeq \frac{T}{q^{3/2}} \left(1 + \frac{t_{\max}}{q^{3/2}} \right) \frac{1}{\mu^2}\,, \tag{27}$$

where estimates (26)-(27) are valid for μ such that

$$\frac{\sqrt{2}\, t_{\max}}{(\mu q)^{3/2}} \ll 1\,.$$

We see that the estimates from Proposition 1 and Theorem 2 decrease faster with distance than the estimate relative to Theorem 1. Moreover, Theorem 2 gives a better result than Proposition 1 if the pulse duration $t_{\max}$ is larger than $q^{3/2}$.

Small distance: $\alpha \ll \mu \ll 1$.

$$\sigma^{\mathrm{Th1}} \simeq \frac{T t_{\max}}{q^3}\, \mathrm{e}^{2 t_{\max}\, q^{-3/2} \mu^{-3}} \frac{2}{\mu^5}\,,$$

$$\sigma^{\mathrm{Prop1}} \simeq \frac{T t_{\max}}{q^3}\, \mathrm{e}^{2 t_{\max}\, q^{-3/2} \mu^{-3}} \frac{1}{\mu^5}\,,$$

$$\sigma^{\mathrm{Th2}} \simeq \frac{T t_{\max}}{q^3}\, \mathrm{e}^{\sqrt{2} t_{\max}\, q^{-3/2} \mu^{-3/2}} \frac{\sqrt{2}}{\mu^{7/2}}\,,$$

where estimates are valid for μ such that

$$\frac{\sqrt{2}\, t_{\max}}{(\mu q)^{3/2}} \gg 1\,.$$

We see that the estimate from Theorem 2 diverges slower than those coming from Theorem 1 and Proposition 1 as μ goes to zero. This is in particular due to the reduction of the exponent in the exponential, from μ^{-3} to $\mu^{-3/2}$. As a consequence, in the two ranges of μ, Theorem 2 provides significantly much better estimates than Theorem 1 and Proposition 1.

One additional advantage with the formulation of Theorem 2 is that it gives separate bounds on the positions and the momenta. These bounds

are represented in Fig. 2 where the positions are rescaled by E_0/ω^2 and the momenta by E_0/ω. Using these scalings, we see that the dominant error as a result of the KH approximation comes from the positions getting close to the ionic core.

3.3. Discussion

As we have seen above, the amended version of Bogolyubov's theorem provides better estimates than the original version. The main reason is that this version is more suited to Hamiltonian systems. Does it allow us to conclude on the validity of the Kramers-Henneberger approximation? As discussed in Sec. 2.3, the answer to the validity of the KH approximation still depends on the values of the parameters and on the region of phase space in which the trajectory evolves. If the trajectory is sufficiently far from the ionic core (at least, at a distance of one quiver radius), the approximation is valid. The validity gets better as the trajectory gets further away from the core. Using a more apt version of the Bogolyubov's averaging theorem adapted to Hamiltonian flows, the validity is confirmed closer to the ionic core than it was the case with the original version of the theorem. However, all versions of the averaging theorem lead to the same conclusions if μ is sufficiently close to zero. This means that when the trajectory is allowed to cross $x = 0$, the bounds between this trajectory and the one of the averaged system become large.

A numerical investigation of H and $\langle H \rangle$ confirms that the two Cauchy problems do not separate exponentially as predicted by averaging theorems. The behavior is much more complex, depending on the region of phase space visited by the trajectories. In order to illustrate this complex behavior a Matlab [27] script `KHBogolyubov.mlx` is available at `github.com/cchandre/KH`. This script computes and displays the trajectories in phase space for both Hamiltonians H and $\langle H \rangle$ starting with the same initial conditions, as well as their distance as a function of time, in order to compare it with the results obtained using Theorem 1, Proposition 1 and Theorem 2. This script also contains the codes to reproduce Figs. 1 and 2.

One of the main arguments of the KH approximation is the possibility to create stable laser-dressed states located near the minimum of the KH potential. However, this minimum is always at a distance smaller than one quiver radius. The estimates obtained by Bogolyubov's theorem or their amended versions do not work in this region of phase space, mostly regardless of the parameters of the laser field. As a consequence, we cannot conclude

on the validity of the KH approximation close to the local minimum of the KH potential which is at the core of the KH atom.

Acknowledgments

CC and JD acknowledge the funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 734557.

References

[1] G. van de Sand, J. M. Rost, Irregular orbits generate higher harmonics, Phys. Rev. Lett. 83 (1999) 524.

[2] P. B. Corkum, Plasma perspective on strong field multiphoton ionization, Phys. Rev. Lett. 71 (1993) 1994.

[3] K. J. Schafer, B. Yang, L. F. DiMauro, K. C. Kulander, Above threshold ionization beyond the high harmonic cutoff, Phys. Rev. Lett. 70 (1993) 1599.

[4] W. Becker, H. Rottke, Many-electron strong-field physics, Cont. Phys. 49 (2008) 199.

[5] R. Grobe, C. K. Law, Stabilization in superintense fields: A classical interpretation, Phys. Rev. A 44 (1991) R4114.

[6] M. Gavrila, Atomic stabilization in superintense laser fields, J. Phys. B: At. Mol. Opt. Phys. 35 (2002) R147.

[7] M. J. Norman, C. Chandre, T. Uzer, P. Wang, Nonlinear dynamics of ionization stabilization of atoms in intense laser fields, Phys. Rev. A 91 (2015) 023406.

[8] H. A. Kramers, Collected Scientific Papers, North-Holland Publishing Company, Amsterdam, 1956.

[9] W. C. Henneberger, Perturbation method for atoms in intense light beams, Phys. Rev. Lett. 21 (1968) 838.

[10] H. Breuer, K. Dietz, M. Holthaus, A remark on the Kramers-Henneberger transformation, Phys. Lett. A 165 (1992) 341.

[11] V. C. Reed, K. Burnett, P. L. Knight, Harmonic generation in the Kramers-Henneberger stabilization regime, Phys. Rev. A 47 (1993) R34.

[12] E. A. Volkova, A. M. Popov, O. V. Smirvova, Stabilization of atoms in a strong field and the Kramers-Henneberger approximation, JETP 79 (1994) 736.

[13] A. M. Popov, O. V. Tikhonova, E. A. Volkova, Applicability of the Kramers-Henneberger approximation in the theory of strong-field ionization, J. Phys. B: At. Mol. Opt. Phys. 32 (1999) 3331.

[14] M. Førre, S. Selstø, J. P. Hansen, L. B. Madsen, Exact nondipole Kramers-Henneberger form of the light-atom hamiltonian: An application to atomic stabilization and photoelectron energy spectra, Phys. Rev. Lett 95 (2005) 043601.

[15] I. A. Ivanov, A. S. Kheifets, On the use of the Kramers–Henneberger Hamiltonian in multi-photon ionization calculations, J. Phys. B: At. Mol. Opt. Phys. 38 (2005) 2245.

[16] P. L. He, Z. H. Zhang, F. He, Young's double-slit interference in a hydrogen atom, Phys. Rev. Lett. 124 (2020) 163201.

[17] F. Morales, M. Richter, S. Patchkovskii, O. V. Smirnova, Imaging the Kramers–Henneberger atom, PNAS 108 (2011) 16906.

[18] M. Richter, F. Patchkovskii, F. Morales, O. Smirnova, M. Ivanov, The role of the Kramers–Henneberger atom in the higher-order Kerr effect, New J. Phys. 15 (2013) 083012.

[19] Q. Wei, P. Wang, S. Kais, D. Herschbach, Pursuit of the Kramers-Henneberger atom, Chem. Phys. Lett. 683 (2017) 240.

[20] O. V. Smirnova, Validity of the Kramers–Henneberger approximation, JETP 90 (2000) 609.

[21] N. Bogolyubov, Y. Mitropol'skii, Asymptotic Methods in the Theory of Non-Linear Oscillations, Gordon and Breach, New York, 1961.

[22] V. F. Zhuravlev, D. M. Klimov, Applied Methods in the Vibration Theory, Nauka, Moscow, 1988, (in Russian).

[23] Y. A. Mitropolsky, N. V. Dao, Applied Asymptotic Methods in Nonlinear Oscillations, Springer Netherlands, 1997.

[24] J. Dubois, S. A. Berman, C. Chandre, T. Uzer, Capturing photoelectron motion with guiding centers, Phys. Rev. Lett. 121 (2018) 113202.

[25] J. Dubois, S. A. Berman, C. Chandre, T. Uzer, Guiding-center motion for electrons in strong laser fields, Phys. Rev. E 98 (2018) 052219.

[26] J. Chandra, P. W. Davis, Linear generalizations of Grönwall's inequality, Proc. Amer. Math. Soc. 60 (1976) 157.

[27] MATLAB, version 9.10.0 (R2021a), The MathWorks Inc., Natick, Massachusetts, 2021.