J. Plasma Phys. (2016), *vol.* 82, 905820304 © Cambridge University Press 2016
doi:10.1017/S0022377816000532

The initial value problem in Lagrangian drift kinetic theory

J. W. Burby†

Courant Institute of Mathematical Sciences, New York, NY 10012, USA

(Received 19 January 2016; revised 20 April 2016; accepted 21 April 2016)

Existing high-order variational drift kinetic theories contain unphysical rapidly varying modes that are not seen at low orders. These unphysical modes, which may be rapidly oscillating, damped or growing, are ushered in by a failure of conventional high-order drift kinetic theory to preserve the structure of its parent model's initial value problem. In short, the (infinite dimensional) system phase space is unphysically enlarged in conventional high-order variational drift kinetic theory. I present an alternative, 'renormalized' variational approach to drift kinetic theory that manifestly respects the parent model's initial value problem. The basic philosophy underlying this alternate approach is that high-order drift kinetic theory ought to be derived by truncating the all-orders system phase-space Lagrangian instead of the usual 'field+particle' Lagrangian. For the sake of clarity, this story is told first through the lens of a finite-dimensional toy model of high-order variational drift kinetics; the analogous full-on drift kinetic story is discussed subsequently. The renormalized drift kinetic system, while variational and just as formally accurate as conventional formulations, does not support the troublesome rapidly varying modes.

1. Introduction

The Lagrangian approach to deriving drift kinetic and gyrokinetic theories (Brizard 2000; Sugama 2000; Brizard & Hahm 2007) has proven to be successful at producing strongly magnetized kinetic plasma models with exact energy and momentum conservation laws (Scott & Smirnov 2010). The philosophy underlying this approach requires all approximations to be implemented at the level of a system Lagrangian. As a result, exact conservation laws are unavoidable, and expressions for the conserved quantities follow systematically from Noether's theorem.

Conservation laws are not everything, however. Another fundamental attribute of a physical model is its phase space (Lee & Wald 1990). Phase space, P, is defined to be any space in which the model is equivalent to a dynamical law for points $p \in P$ of the form $\mathrm{d}p/\mathrm{d}t = V(p, t)$, where V is a time-dependent vector field on P. Before any mention can be made of a dynamical system's conservation laws, the 'arena' where the dynamics take place must be identified; phase space is precisely this arena. So while it is natural to demand that kinetic plasma models have correct conservation laws, it is equally natural to require that these models have correct phase spaces.

† Email address for correspondence: Joshua.Burby@cims.nyu.edu

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

The question I will investigate in this paper concerns the fate of phase space in Lagrangian drift kinetic and gyrokinetic theories. While the Lagrangian method gets the conservation laws right, does it produce a model with the correct phase space? Equivalently, do these theories require too many, too few, or just the right number of initial conditions?

Perhaps surprisingly, the answer to this question depends on the order of accuracy (in $\epsilon = \rho/L$) of the gyrokinetic Lagrangian. At very low orders, the question is not interesting; the answer is 'just the right number.' In high-flow drift kinetics, 'very low' means retaining terms in the guiding centre Hamiltonian that are zeroth order in magnetic non-uniformity. In low-flow drift kinetics and gyrokinetics, 'very low' means retaining terms in the Hamiltonian that are second order in magnetic non-uniformity. On the other hand, at higher orders than these specified, the question is much more interesting. In particular, high-flow drift kinetics with a first-order (in magnetic nonuniformity) guiding centre Hamiltonian requires special attention.

In § 3, I will show that Miyato *et al.* (2009)'s variational high-flow first-order electrostatic drift kinetics, which I will refer to as MSST drift kinetics (the acronym corresponds to the names of the authors of (Miyato *et al.* 2009)), has an unphysically large phase space. Where one should only have to specify F at $t = 0$ to determine the subsequent evolution of the drift kinetic system, this theory requires φ to be supplied independently at $t = 0$ as well[1]. This problem is caused by the appearance of a time derivative of the electrostatic potential in the drift kinetic Poisson equation and guiding centre equations of motion. As a consequence, the drift kinetic system exhibits certain malignant features. I argue using linearized equations that MSST drift kinetics supports unphysical rapidly varying modes, akin to the 'runaway' solutions allowed by the Lorentz–Abraham–Dirac equation describing radiation reaction. These modes are rapidly varying in the sense that their frequency increases as ϵ decreases. Moreover, their presence is entirely a consequence of requiring too many initial conditions; in lower-order drift kinetics, where the size of phase space is correct, the modes do not appear.

Fortunately, there is a way to rescue the Lagrangian drift kinetic method from these doldrums. In § 4, I present a systematic 'renormalized' Lagrangian approach to MSST drift kinetics that manifestly retains the physical phase space. The basic idea underlying this reformulation (and slight modification) of MSST drift kinetics is that the system phase-space Lagrangian for all-orders drift kinetics is a better starting point for performing asymptotics than the 'particle phase-space Lagrangian + electromagnetic Lagrangian,' which is the starting point in all previous treatments of variational drift kinetics and gyrokinetics. This perspective enables the formulation of a variational version of the following basic procedure for eliminating extra time derivatives of the potential. Take a time derivative of the drift kinetic Poisson equation to find an approximate expression for the time derivative of the potential in terms of the distribution function and the potential. Then substitute this expression back into the drift kinetic Poisson equation and guiding centre equations of motion.

One great benefit of formulating a variational version of this procedure is that all exact conservation laws of the drift kinetic system are preserved automatically, which would likely not happen if the non-variational version of the technique were applied to solve for the time derivatives. Another is that the numerous specialized

[1]It is important to realize that in Miyato *et al.* (2009) a subsidiary expansion is invoked in order to eliminate the time derivative of the potential that ultimately leads to this bad behaviour. The MSST drift kinetic system studied in this article is the drift kinetic system presented in Miyato *et al.* (2009) before the subsidiary ordering is introduced.

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

tools developed for the analysis of infinite dimensional Lagrangian and Hamiltonian systems may be applied to renormalized MSST drift kinetic theory; see Morrison & Kotschenreuther (1990) for free energy principles, Craig & Wayne (1994) for existence of periodic solutions, Vanneste & Wirosoetisno (2006) for adiabatic invariant theory, Wayne (1990) for KAM theory, and Bourgain (1996) for statistical mechanics of Hamiltonian systems. In short, the method is structure-preserving.

Because the derivation of the renormalized Lagrangian involves some relatively advanced techniques from the calculus of variations, I have devoted § 2 to a description of the necessary manipulations in the context of a finite-dimensional toy model of drift kinetics. This toy model illustrates why standard Lagrangian drift kinetics does violence to the drift kinetic phase space without introducing the infinite-dimensional machinery used in drift kinetic field theory. As such, § 2 can be viewed as a simplified version of the entire article.

While this article focuses on a particular drift kinetic model, the ideas it introduces and the techniques it employs are important in a broader context. Most practically, the difficulties introduced by extra time derivatives of the electromagnetic field, and the method to circumvent these difficulties within the Lagrangian framework, are also highly relevant to low-flow gyrokinetic models. This point is driven home by the recent toroidal momentum transport controversy. See Krommes & Hammett (2013) for a summary and critical assessment of the different sides of the debate, and Parra & Catto (2008) for the paper that initiated the discussion. In Parra *et al.* (2012), the authors argue that certain high-order terms in the gyrocentre Hamiltonian (the $O(\rho^{*3}T_e)$ terms) should be retained in a full-f gyrokinetic approach in order to accurately model the transport of toroidal momentum in the low-flow regime on transport time scales. These high-order terms include time derivatives of the electromagnetic field. Therefore, whenever the high-order terms highlighted in Parra *et al.* (2012) are included in a gyrokinetic model (presumably when doing low-flow modelling in the core), the results in the present article will be essential to understand. Moreover, it is worth mentioning that without the method proposed here for dealing with the extra time derivatives of the electromagnetic field, the mathematical and physical soundness of the variational gyrokinetic and drift kinetic approximation would be in question. If the behaviour of high-order gyrokinetics differed significantly from that of low-order gyrokinetics (for instance, due to the presence of unphysical rapidly varying modes), a dark shadow would be cast on the results predicted by the low-order theory.

Speaking more generally than gyrokinetic and drift kinetic models, the methods used in this article will likely prove useful for developing reduced variational models for a wide variety of plasma systems. It is not uncommon to encounter plasma models with high-order time derivatives multiplied by small parameters. The highest-order time derivative tends to set the number of required initial conditions in a model, and a small parameter multiplying this derivative generally indicates the presence of rapidly varying dynamical behaviour. While in the drift kinetic and gyrokinetic systems considered here, the rapidly varying dynamical behaviour is unphysical, this is not always the case. For instance, the Vlasov–Maxwell system can be expressed as the Vlasov–Darwin system plus small higher-order time derivatives. Nevertheless, it is always interesting to study any possible slow dynamical behaviour that such a singularly perturbed plasma model might allow for. In general, slow dynamics will occur when initial conditions are sampled from the so-called slow manifold (Jones 1995). The methods employed in this article can be used to find a variational formulation for dynamics on any Lagrangian system's slow manifold. I am currently

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

using them to identify a variational formulation of dynamics on a slow manifold contained in the Vlasov–Maxwell system. Another interesting application would be identifying a variational principle for the generalized reduced MHD system presented in Kruger, Hegna & Callen (1998).

2. Two-oscillator drift kinetic theory

2.1. *Introducing the all-orders theory*

In this section I will analyse the initial value problem for a toy model I will call two-oscillator drift kinetics. The motivation for studying this relatively simple model is that its initial value problem is structurally very similar to that of drift kinetics. In this first subsection I will give a derivation of the all-orders version of the model, which starts from the following system of two coupled one-dimensional oscillators,

$$\ddot{q} = -\omega_1^2 q - 2gq\varphi \tag{2.1}$$

$$\ddot{\varphi} = -\omega_2^2 \varphi - gq^2. \tag{2.2}$$

I will refer to the dynamical system defined by (2.1) and (2.2) as two-oscillator kinetics. The quantities $q, \varphi \in \mathbb{R}$ are the oscillator positions. As the notation suggests, q-space should be thought of as an analogue of the single-particle configuration space. Similarly, φ should be thought of as the electromagnetic field. The parameters g, ω_1, ω_2 represent the coupling constant, the characteristic frequency for particles and the characteristic frequency for the fields, respectively. This system can be derived from the Lagrangian

$$L(q, p, \varphi, \dot{q}, \dot{p}, \dot{\varphi}) = L_p(q, p, \dot{q}, \dot{p}) + L_{int}(q, \varphi) + L_f(\varphi, \dot{\varphi}), \tag{2.3}$$

where

$$L_p = p\,\dot{q} - \left(\tfrac{1}{2}p^2 + \tfrac{1}{2}\omega_1^2 q^2\right) \tag{2.4}$$

$$L_f = \tfrac{1}{2}\dot{\varphi}^2 - \tfrac{1}{2}\omega_2^2 \varphi^2 \tag{2.5}$$

$$L_{int} = -gq^2\varphi, \tag{2.6}$$

and p denotes the particle canonical momentum. Note that the free particle Lagrangian L_p is a so-called 'phase-space Lagrangian.' In place of L_p, the 'configuration-space Lagrangian,' $\tilde{L}_p = \dot{q}^2/2 - \omega_1^2 q^2/2$, could have been substituted, but the phase-space Lagrangian allows greater freedom in applying coordinate transformations to the particle variables. It is for this same reason that derivations of drift kinetics necessarily employ a phase-space Lagrangian for particles. The equations of motion corresponding to this system Lagrangian are found by varying the action $S = \int_{t_1}^{t_2} L\,\mathrm{d}t$ while holding the values of q, p, and φ fixed at t_1 and t_2. In particular, extremizing the action leads to the Euler–Lagrange equations,

$$\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} = 0 \tag{2.7}$$

$$\frac{\partial L}{\partial p} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{p}} = 0 \tag{2.8}$$

$$\frac{\partial L}{\partial \varphi} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\varphi}} = 0, \tag{2.9}$$

which are readily seen to reproduce (2.1) and (2.2).

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

All-orders two-oscillator drift kinetics is obtained from this 'particle-space' model by passing from particle coordinates (q, p) to 'guiding centre coordinates' (Q, P). Just as in drift kinetic theory, the guiding centre coordinates are related to the particle coordinates by a transformation of the particle phase space that depends parametrically on the first N time derivatives of the field φ, where N is some positive integer. Let $j\varphi = (\varphi, \partial_t\varphi, \ldots, \partial_t^N\varphi)$ and $T_{j\varphi} : \mathbb{R}^2 \to \mathbb{R}^2$ be the gyrocentre coordinate transformation. By definition,

$$(Q, P) = T_{j\varphi}(q, p) \equiv T(q, p; j\varphi). \tag{2.10}$$

Note that this transformation can be thought of as passing into a moving frame on the particle phase space; the frame moves as the field φ evolves.

For the sake of concreteness and simplicity, assume $T_{j\varphi}$ is given by

$$T_{j\varphi}(q, p) = (q + \epsilon\ddot{\varphi}, p), \tag{2.11}$$

where ϵ is a constant with dimensions of inverse frequency squared. The inverse of $T_{j\varphi}$ is therefore

$$T_{j\varphi}^{-1}(Q, P) = (Q - \epsilon\ddot{\varphi}, P). \tag{2.12}$$

Using these explicit expressions for the guiding centre transformation, I will now determine the differential equations governing all-orders two-oscillator drift kinetics.

Suppose that q, p, φ satisfy the dynamical equations of two-oscillator kinetics. I will find differential equations governing the quantities $Q = q + \epsilon\ddot{\varphi}$, $P = p$ and φ. Using the field equation for two-oscillator kinetics, it is straightforward to see that Q is related to q and φ by

$$Q = q - \epsilon g q^2 - \epsilon\omega_2^2\varphi, \tag{2.13}$$

which also gives a relationship between the time derivatives of Q, q, and φ according to

$$\dot{Q} = (1 - 2\epsilon g q)\dot{q} - \epsilon\omega_2^2\dot{\varphi}. \tag{2.14}$$

Therefore whenever $q < 1/2\epsilon g$,

$$\begin{aligned} q &= \frac{1}{2\epsilon g}\left(1 - \sqrt{1 - 4\epsilon g Q - 4\epsilon^2\omega_2^2 g\varphi}\right) \\ &= Q + \epsilon(gQ^2 + \omega_2^2\varphi) + O(\epsilon^2) \\ &= \hat{q}_-(Q, \varphi), \end{aligned} \tag{2.15}$$

and whenever $q > 1/2\epsilon g$,

$$\begin{aligned} q &= \frac{1}{2\epsilon g}\left(1 + \sqrt{1 - 4\epsilon g Q - 4\epsilon^2\omega_2^2 g\varphi}\right) \\ &= \frac{1}{\epsilon g} - Q - \epsilon(gQ^2 + \omega_2^2\varphi) + O(\epsilon^2) \\ &= \hat{q}_+(Q, \varphi). \end{aligned} \tag{2.16}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

Now consider an interval of time when $q < 1/2\epsilon g$. Using the previous expressions, $P = p$ and $\dot{q} = p$, it is straightforward to verify that $\dot{Q}$ and $\dot{P}$ are given by

$$\dot{P} = -(\omega_1^2 + 2g\varphi)\hat{q}_-(Q, \varphi) \tag{2.17}$$

$$\dot{Q} = (1 - 2\epsilon g\hat{q}_-(Q, \varphi))P - \epsilon\omega_2^2\dot{\varphi} \tag{2.18}$$

$$\ddot{\varphi} = -\omega_2^2\varphi - g\hat{q}_-^2(Q, \varphi). \tag{2.19}$$

This system of equations governs all-orders two-oscillator drift kinetics when $q < 1/2\epsilon g$. When $q > 1/2\epsilon g$, the governing equations are instead

$$\dot{P} = -(\omega_1^2 + 2g\varphi)\hat{q}_+(Q, \varphi) \tag{2.20}$$

$$\dot{Q} = (1 - 2\epsilon g\hat{q}_+(Q, \varphi))P - \epsilon\omega_2^2\dot{\varphi} \tag{2.21}$$

$$\ddot{\varphi} = -\omega_2^2\varphi - g\hat{q}_+^2(Q, \varphi). \tag{2.22}$$

Perplexingly, the preceding analysis has identified two sets of evolution equations for all-orders two-oscillator drift kinetics. Equivalently, the analysis produced a multi-valued vector field on $(P, Q, \varphi, \dot{\varphi})$-space. This multi-valuedness does not signal a breakdown of determinism. Instead it indicates that $(P, Q, \varphi, \dot{\varphi})$-space is two-sheeted, which is in stark contrast to the single-sheeted $(p, q, \varphi, \dot{\varphi})$-space. The cut in the sheet occurs along the line in the (Q, φ)-plane given by

$$D = 1 - 4\epsilon gQ - 4\epsilon^2\omega_2^2 g\varphi = 0. \tag{2.23}$$

On the 'top' sheet $\hat{q}_+$ is used in the equations of motion, while on the 'bottom' sheet $\hat{q}_-$ is used. When the system reaches the cut, one of two things happens. If $P = p = \dot{q} > 0$, then the subsequent evolution of the system occurs on the top sheet. If $P < 0$, then the subsequent evolution of the system occurs on the bottom sheet. With this knowledge of the sheet structure of $(P, Q, \varphi, \dot{\varphi})$-space together with our multi-valued dynamical vector field, the evolution equations for all-orders two-oscillator drift kinetics are completely specified.

I will conclude this subsection by presenting the Lagrangian formulation of the all-orders two-oscillator drift kinetic system. The Lagrangian is found by substituting the guiding centre transformation into the two-oscillator kinetics Lagrangian, which gives

$$\begin{aligned}\bar{L}(Q, P, \varphi, \dot{Q}, \dot{P}, \dot{\varphi}, \ddot{\varphi}, \dddot{\varphi}) \;=\; & L(Q, P, \varphi, \dot{Q}, \dot{P}, \dot{\varphi}) + \epsilon(Q\ddot{\varphi}(2g\varphi + \omega_1^2) - P\dddot{\varphi}) \\ & - \frac{\epsilon^2}{2}\ddot{\varphi}^2(2g\varphi + \omega_1^2).\end{aligned} \tag{2.24}$$

Because this Lagrangian depends on the third time derivative of φ, its associated Euler–Lagrange equations are given by

$$\frac{\partial\bar{L}}{\partial Q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\bar{L}}{\partial\dot{Q}} = 0 \tag{2.25}$$

$$\frac{\partial\bar{L}}{\partial P} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\bar{L}}{\partial\dot{P}} = 0 \tag{2.26}$$

$$\frac{\partial\bar{L}}{\partial\varphi} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\bar{L}}{\partial\dot{\varphi}} + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial\bar{L}}{\partial\ddot{\varphi}} - \frac{\mathrm{d}^3}{\mathrm{d}t^3}\frac{\partial\bar{L}}{\partial\dddot{\varphi}} = 0. \tag{2.27}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

It is straightforward to verify that these equations reduce to

$$-\dot{P} - (Q - \epsilon\ddot{\varphi})(2g\varphi + \omega_1^2) = 0 \tag{2.28}$$

$$-P + \dot{Q} - \epsilon\dddot{\varphi} = 0 \tag{2.29}$$

$$\begin{aligned}-\ddot{\varphi} - \omega_2^2\varphi - gQ^2 + \epsilon(\ddot{P} + 4g\dot{Q}\dot{\varphi} + 4gQ\ddot{\varphi} + \ddot{Q}(2g\varphi + \omega_1^2)) \\ - \epsilon^2(\dddot{\varphi}(2g\varphi + \omega_1^2) + 3g\ddot{\varphi}^2 + 4g\dot{\varphi}\dddot{\varphi}) = 0.\end{aligned} \tag{2.30}$$

Using (2.28), $\ddot{P}$ can be expressed in terms of Q, φ and time derivatives of φ. When this expression is substituted into (2.30), all third- and fourth-time derivatives of φ exactly cancel, leaving a quadratic equation for $\ddot{\varphi}$. The two roots of this quadratic equation correspond to the two sheets identified earlier, and each expresses $\ddot{\varphi}$ in terms of Q and φ. When either of the expressions for $\ddot{\varphi}$ are substituted into (2.28) and (2.29), evolution equations for P and Q emerge. The resulting system of equations is equivalent to the equations just derived by manipulating the two-oscillator kinetic equations directly.

2.2. *Initial value problem for the all-orders model*

Because I was able to write the differential equations governing all-orders two-oscillator drift kinetics explicitly, it is not difficult to see that the phase space for all-orders two-oscillator drift kinetics has the same dimension as the underlying particle-space theory, two-oscillator kinetics. In each case, four numbers must be specified at $t = 0$ in order to uniquely determine a solution. Interestingly, in passing from particle-space to guiding centre phase space, the phase space for the field-particle system becomes folded, i.e. multi-sheeted.

While it is not simple to gain a complete understanding of the dynamical behaviour predicted by all-orders two-oscillator drift kinetics, it is straightforward to determine the characteristic frequencies of oscillations about the bottom-sheet equilibrium $(Q, P, \varphi, \dot{\varphi}) = 0$. Linearizing equations (2.17)–(2.19) about this equilibrium gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta P = -\omega_1^2(\delta Q + \epsilon\omega_2^2\delta\varphi) \tag{2.31}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta Q = \delta P - \epsilon\omega_2^2\delta\dot{\varphi} \tag{2.32}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta\varphi = \delta\dot{\varphi} \tag{2.33}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta\dot{\varphi} = -\omega_2^2\delta\varphi, \tag{2.34}$$

which can be written as $\dot{z} = Az$, where $z = (\delta P, \delta Q, \delta\varphi, \delta\dot{\varphi})$ and A is a 4×4 matrix. The spectrum of A is given by $\{\mathrm{i}\omega_1, -\mathrm{i}\omega_1, \mathrm{i}\omega_2, -\mathrm{i}\omega_2\}$. This indicates that the normal modes of all-orders two-oscillator drift kinetics linearized about the origin are purely oscillatory.

2.3. *Two oscillator drift kinetics and its initial value problem*

Because I did not make any approximations in passing from the underlying particle-space model to the guiding centre model, all-orders two-oscillator drift kinetics is an exact theory. The same cannot be said of ordinary drift kinetic theory. In that theory, the coordinate transformation that relates guiding centre coordinates to particle

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

coordinates is only known as an asymptotic series in $\epsilon = \rho/L$. As a result, the system Lagrangian for ordinary drift kinetics is not known exactly, but only asymptotically. In particular, the drift kinetic Lagrangian is always expanded in powers of ϵ and then truncated at some order.

For the sake of examining the consequences of this incomplete knowledge of the system Lagrangian in drift kinetic theory, let us now expand the two-oscillator drift kinetic Lagrangian in powers of ϵ and then drop terms that are $O(\epsilon^2)$ or higher, and then re-do the initial value analysis. I will refer to the dynamical system defined by the truncated Lagrangian simply as two-oscillator drift kinetics. The truncated Lagrangian is given by

$$\bar{L}(Q, P, \varphi, \dot{Q}, \dot{P}, \dot{\varphi}, \ddot{\varphi}) = L(Q, P, \varphi, \dot{Q}, \dot{P}, \dot{\varphi}) + \epsilon(Q\ddot{\varphi}(2g\varphi + \omega_1^2) - P\dddot{\varphi}), \tag{2.35}$$

which again depends on the third time derivative of φ. The Euler–Lagrange equations are now

$$-\dot{P} - (Q - \epsilon\ddot{\varphi})(2g\varphi + \omega_1^2) = 0 \tag{2.36}$$

$$-P + \dot{Q} - \epsilon\dddot{\varphi} = 0 \tag{2.37}$$

$$-\ddot{\varphi} - \omega_2^2\varphi - gQ^2 + \epsilon(\ddot{P} + 4g\dot{Q}\dot{\varphi} + 4gQ\ddot{\varphi} + \ddot{Q}(2g\varphi + \omega_1^2)) = 0. \tag{2.38}$$

As before, (2.36) can be used to derive an expression for $\ddot{P}$ that can be substituted into (2.38) in order to relate Q, φ and the time derivatives of φ. In contrast to the all-orders theory, the third- and fourth-derivatives of φ do not cancel, and the following evolution equations for two-oscillator drift kinetics emerge:

$$\dot{P} = -(Q - \epsilon\ddot{\varphi})(2g\varphi + \omega_1^2) \tag{2.39}$$

$$\dot{Q} = P + \epsilon\dddot{\varphi} \tag{2.40}$$

$$\ddddot{\varphi} = \frac{1}{\epsilon^2(2g\varphi + \omega_1^2)}(\ddot{\varphi} + gQ^2 + \omega_2^2\varphi - 2\epsilon gQ\ddot{\varphi} - 2\epsilon^2 g(\ddot{\varphi}^2 + 2\dot{\varphi}\dddot{\varphi})). \tag{2.41}$$

Evidently the phase space for two-oscillator drift kinetic theory is $(P, Q, \varphi, \dot{\varphi}, \ddot{\varphi}, \dddot{\varphi})$-space. This is clearly in disagreement with the all-orders theory, which has a four-dimensional two-sheeted phase space. The Lagrangian approximation method has failed in a qualitative manner! This failure can also be detected by examining the behaviour of two-oscillator drift kinetics linearized about the equilibrium $(P, Q, \varphi, \dot{\varphi}, \ddot{\varphi}, \dddot{\varphi}) = 0$. The linearized equations are given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta P = -\omega_1^2(\delta Q - \epsilon\delta\ddot{\varphi}) \tag{2.42}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta Q = \delta P + \epsilon\delta\dddot{\varphi} \tag{2.43}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta\varphi = \delta\dot{\varphi} \tag{2.44}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta\dot{\varphi} = \delta\ddot{\varphi} \tag{2.45}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta\ddot{\varphi} = \delta\dddot{\varphi} \tag{2.46}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta\dddot{\varphi} = \frac{1}{\epsilon^2\omega_1^2}(\delta\ddot{\varphi} + \omega_2^2\delta\varphi), \tag{2.47}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

which can be written as $\dot{z} = Az$, where $z = (\delta P, \delta Q, \delta\varphi, \delta\dot{\varphi}, \delta\ddot{\varphi}, \delta\dddot{\varphi})$ and A is a 6×6 matrix. The spectrum of A is given by $\{\mathrm{i}\omega_1, -\mathrm{i}\omega_1, i\tilde{\omega}_2, -i\tilde{\omega}_2, \gamma_\times, -\gamma_\times\}$, where

$$\tilde{\omega}_2 = \omega_2 \left(1 - \tfrac{1}{2}\omega_1^2\omega_2^2\epsilon^2\right) + O(\epsilon^3) \tag{2.48}$$

$$\gamma_\times = \frac{1}{\epsilon\omega_1}\left(1 + \tfrac{1}{2}\omega_1^2\omega_2^2\epsilon^2\right) + O(\epsilon^3). \tag{2.49}$$

In contrast to the all-orders theory, linearized two-oscillator drift kinetics therefore has rapidly growing normal modes with growth rate $\gamma_\times$. The reason for the presence of these troublesome modes is the increased dimensionality of the two-oscillator drift kinetic phase space; it has nothing to do with the sheet structure of this space.

As I will show in more detail in the next section, conventional Lagrangian drift kinetic theory suffers from problems similar to those that I have highlighted in the context of two-oscillator drift kinetics. In both systems, the root cause of the trouble is the combination of (1) employing a coordinate change on the single-particle phase space that depends on time derivatives of the electromagnetic fields and (2) truncating a 'field+particle' system Lagrangian.

2.4. *Rescuing phase space*

This article is not meant to paint a depressing picture of Lagrangian drift kinetics. While the flaw with the conventional Lagrangian approach identified in the previous subsection is important to be aware of, it is equally important to know that the flaw is curable. In this subsection I will describe the cure in the context of two-oscillator drift kinetics.

In order to avoid the possibility of destroying the physical structure of phase space without resorting to the all-orders theory, I advocate truncating the phase-space Lagrangian for the entire two-oscillator drift kinetic system. See appendix A for a discussion of the precise meaning of the term 'phase-space Lagrangian.' I will explain why this is preferable to the Lagrangian approach described in the previous subsection by first showing that it works and then giving a brief abstract discussion.

First I will derive the phase-space Lagrangian for all-orders two-oscillator drift kinetics. See appendix A, which is a partial generalization of a discussion found in Marsden, Patrick & Shkoller (1998), for a rigorous justification of the procedure I will use to do this. In the derivation I will restrict attention to the bottom sheet (corresponding to $\hat{q}_-$); the top sheet can be treated in the same manner. Consider the free-endpoint variation of the all-orders two-oscillator drift kinetic Lagrangian. A simple calculation shows

$$\begin{aligned}\delta\int_{t_1}^{t_2}\bar{L}\,\mathrm{d}t \;=\; &\int_{t_1}^{t_2}\left(\frac{\partial\bar{L}}{\partial Q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\bar{L}}{\partial\dot{Q}}\right)\delta Q\,\mathrm{d}t\\ &+\int_{t_1}^{t_2}\left(\frac{\partial\bar{L}}{\partial P} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\bar{L}}{\partial\dot{P}}\right)\delta P\,\mathrm{d}t\\ &+\int_{t_1}^{t_2}\left(\frac{\partial\bar{L}}{\partial\varphi} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\bar{L}}{\partial\dot{\varphi}} + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial\bar{L}}{\partial\ddot{\varphi}} - \frac{\mathrm{d}^3}{\mathrm{d}t^3}\frac{\partial\bar{L}}{\partial\dddot{\varphi}}\right)\delta\varphi\,\mathrm{d}t\\ &+[\vartheta_Q\delta Q + \vartheta_\varphi\delta\varphi + \vartheta_{\dot{\varphi}}\delta\dot{\varphi} + \vartheta_{\ddot{\varphi}}\delta\ddot{\varphi}]_{t_1}^{t_2},\end{aligned} \tag{2.50}$$

where

$$\vartheta_Q = \frac{\partial\bar{L}}{\partial\dot{Q}} \tag{2.51}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

$$\vartheta_{\varphi} = \frac{\partial \bar{L}}{\partial \dot{\varphi}} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \bar{L}}{\partial \ddot{\varphi}} + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial \bar{L}}{\partial \dddot{\varphi}} \tag{2.52}$$

$$\vartheta_{\dot{\varphi}} = \frac{\partial \bar{L}}{\partial \ddot{\varphi}} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \bar{L}}{\partial \dddot{\varphi}} \tag{2.53}$$

$$\vartheta_{\ddot{\varphi}} = \frac{\partial \bar{L}}{\partial \dddot{\varphi}}. \tag{2.54}$$

There is not a contribution in the boundary term on the last line of (2.50) proportional to δP because $\bar{L}$ does not depend on any time derivatives of P. I will refer to the 1-form on $(Q, P, \dot{Q}, \dot{P}, \varphi, \dot{\varphi}, \ddot{\varphi}, \dddot{\varphi})$-space given by

$$\vartheta^{\infty} = \vartheta_Q \,\mathrm{d}Q + \vartheta_{\varphi}\,\mathrm{d}\varphi + \vartheta_{\dot{\varphi}}\,\mathrm{d}\dot{\varphi} + \vartheta_{\ddot{\varphi}}\,\mathrm{d}\ddot{\varphi} \tag{2.55}$$

as the bare all-orders Lagrange 1-form. As described in Marsden *et al.* (1998), if the Lagrangian $\bar{L}$ for all-orders two-oscillator drift kinetics were non-degenerate, then ϑ^{∞} would give the 'symplectic part' of the system phase space Lagrangian on $(Q, P, \dot{Q}, \dot{P}, \varphi, \dot{\varphi}, \ddot{\varphi}, \dddot{\varphi})$-space. However, because all higher time derivatives of φ than $\ddot{\varphi}$ cancel out of the Euler–Lagrange equations, $\bar{L}$ is degenerate, and I cannot deduce the phase-space Lagrangian so easily. Moreover, it is a good thing that this technique does not work to determine the symplectic part of the system phase-space Lagrangian because $(Q, P, \dot{Q}, \dot{P}, \varphi, \dot{\varphi}, \ddot{\varphi}, \dddot{\varphi})$-space is not the correct phase space for all-orders two-oscillator drift kinetics! The correct symplectic part of the phase-space Lagrangian is given by the dressed Lagrange 1-form $\tilde{\vartheta}^{\infty}$. Here 'dressing' means using the all-orders two-oscillator drift kinetic equations of motion, (2.17)–(2.19), to write the bare Lagrange 1-form in terms of only the true phase-space variables $(Q, P, \varphi, \dot{\varphi})$.

This calculation proceeds as follows. Formally, the dressed Lagrange 1-form is given by

$$\tilde{\vartheta}^{\infty} = \tilde{\vartheta}_Q\,\tilde{\mathrm{d}}Q + \tilde{\vartheta}_{\varphi}\,\tilde{\mathrm{d}}\varphi + \tilde{\vartheta}_{\dot{\varphi}}\,\tilde{\mathrm{d}}\dot{\varphi} + \tilde{\vartheta}_{\ddot{\varphi}}\,\tilde{\mathrm{d}}\ddot{\varphi}, \tag{2.56}$$

where the presence of a tilde indicates a dressed quantity. Because the differentials $\mathrm{d}Q$, $\mathrm{d}\varphi$, $\mathrm{d}\dot{\varphi}$ are already expressed as 1-forms on $(Q, P, \varphi, \dot{\varphi})$-space, dressing does nothing to them, i.e.

$$\tilde{\mathrm{d}}Q = \mathrm{d}Q \tag{2.57}$$

$$\tilde{\mathrm{d}}\varphi = \mathrm{d}\varphi \tag{2.58}$$

$$\tilde{\mathrm{d}}\dot{\varphi} = \mathrm{d}\dot{\varphi}. \tag{2.59}$$

On the other hand, the differential $\mathrm{d}\ddot{\varphi}$ does not live on $(Q, P, \varphi, \dot{\varphi})$-space, and so applying the dressing operation to it is non-trivial. By (2.19), the dressed differential of $\ddot{\varphi}$ can be expressed as

$$\begin{aligned}
\tilde{\mathrm{d}}\ddot{\varphi} &= -\omega_2^2\,\mathrm{d}\varphi - 2g\hat{q}_-\mathrm{d}\hat{q}_- \\
&= -\omega_2^2\,\mathrm{d}\varphi - \frac{2g\hat{q}_-}{\sqrt{D}}(\mathrm{d}Q + \epsilon\omega_2^2\,\mathrm{d}\varphi) \\
&= -\omega_2^2\left(1 + \frac{1-\sqrt{D}}{\sqrt{D}}\right)\mathrm{d}\varphi - \frac{1-\sqrt{D}}{\epsilon\sqrt{D}}\,\mathrm{d}Q \\
&= -\frac{\omega_2^2}{\sqrt{D}}\,\mathrm{d}\varphi - \frac{1-\sqrt{D}}{\epsilon\sqrt{D}}\,\mathrm{d}Q,
\end{aligned} \tag{2.60}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

which is a 1-form on $(Q, P, \varphi, \dot{\varphi})$ space, as required. Here I have used the identity $2\epsilon g\hat{q}_- = 1 - \sqrt{D}$. To dress the components of ϑ^∞, the equations of motion must be used in an analogous fashion. Two of the components require no use of the equations of motion whatsoever, namely

$$\tilde{\vartheta}_Q = P \tag{2.61}$$

$$\tilde{\vartheta}_{\dot{\varphi}} = -\epsilon P. \tag{2.62}$$

Dressing the remaining two components is greatly simplified by making use of the identity

$$Q - \hat{q}_- = \epsilon\ddot{\varphi}, \tag{2.63}$$

which follows from a second useful identity,

$$1 - 4\epsilon gQ - 4\epsilon^2\omega_2^2 g\varphi = D = 1 - 4\epsilon g\hat{q}_- + 4\epsilon^2 g^2\hat{q}_-^2, \tag{2.64}$$

and (2.19). Applying these identities to $\tilde{\vartheta}_\varphi$ gives

$$\begin{aligned}
\tilde{\vartheta}_\varphi &= \frac{\partial\bar{L}}{\partial\dot{\varphi}} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\bar{L}}{\partial\ddot{\varphi}} + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial\bar{L}}{\partial\dddot{\varphi}} \\
&= \dot{\varphi} - \frac{\mathrm{d}}{\mathrm{d}t}(\epsilon Q(2g\varphi + \omega_1^2) - \epsilon^2\ddot{\varphi}(2g\varphi + \omega_1^2)) + \frac{\mathrm{d}}{\mathrm{d}t}(-\epsilon\dot{P}) \\
&= \dot{\varphi} - \frac{\mathrm{d}}{\mathrm{d}t}(\epsilon Q(2g\varphi + \omega_1^2) - \epsilon^2\ddot{\varphi}(2g\varphi + \omega_1^2)) + \frac{\mathrm{d}}{\mathrm{d}t}(\epsilon\hat{q}_-(2g\varphi + \omega_1^2)) \\
&= \dot{\varphi} - \frac{\mathrm{d}}{\mathrm{d}t}((Q - \hat{q}_- - \epsilon\ddot{\varphi})\epsilon(2g\varphi + \omega_1^2)) \\
&= \dot{\varphi}.
\end{aligned} \tag{2.65}$$

Similarly, $\tilde{\vartheta}_{\dot{\varphi}}$ is given by

$$\begin{aligned}
\tilde{\vartheta}_{\dot{\varphi}} &= \frac{\partial\bar{L}}{\partial\ddot{\varphi}} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\bar{L}}{\partial\dddot{\varphi}} \\
&= (\epsilon Q(2g\varphi + \omega_1^2) - \epsilon^2\ddot{\varphi}(2g\varphi + \omega_1^2)) - (-\epsilon\dot{P}) \\
&= (Q - \hat{q}_- - \epsilon\ddot{\varphi})\epsilon(2g\varphi + \omega_1^2) \\
&= 0.
\end{aligned} \tag{2.66}$$

It follows that the dressed Lagrange 1-form is given by

$$\begin{aligned}
\tilde{\vartheta}^\infty &= P\,\mathrm{d}Q + \dot{\varphi}\,\mathrm{d}\varphi + 0\,\mathrm{d}\dot{\varphi} + \epsilon P\left(\frac{\omega_2^2}{\sqrt{D}}\,\mathrm{d}\varphi + \frac{1 - \sqrt{D}}{\epsilon\sqrt{D}}\,\mathrm{d}Q\right) \\
&= P\left(1 + \frac{1 - \sqrt{D}}{\sqrt{D}}\right)\mathrm{d}Q + \left(\dot{\varphi} + \frac{\epsilon P\omega_2^2}{\sqrt{D}}\right)\mathrm{d}\varphi \\
&= \frac{P}{\sqrt{D(Q,\varphi)}}\,\mathrm{d}Q + \left(\dot{\varphi} + \frac{\epsilon\omega_2^2 P}{\sqrt{D(Q,\varphi)}}\right)\mathrm{d}\varphi.
\end{aligned} \tag{2.67}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

The phase-space action for all-orders two-oscillator drift kinetics is therefore

$$S(\gamma) = \int_\gamma \tilde{\vartheta}^\infty - \int_{t_1}^{t_2} \tilde{H}^\infty(\gamma(t))\,\mathrm{d}t, \tag{2.68}$$

where γ is the curve in $(Q, P, \varphi, \dot{\varphi})$-space parameterized by the time interval $[t_1, t_2]$ and $\tilde{H}^\infty$ is the dressed Hamiltonian. The dressed Hamiltonian is obtained from the usual 'bare' Hamiltonian,

$$\begin{aligned} H(Q, P, \dot{Q}, \dot{P}, \varphi, \dot{\varphi}, \ddot{\varphi}, \dddot{\varphi}) \;=\; & \frac{\partial \bar{L}}{\partial \dot{Q}}\dot{Q} + \frac{\partial \bar{L}}{\partial \dot{P}}\dot{P} + \frac{\partial \bar{L}}{\partial \dddot{\varphi}}\dddot{\varphi} + \left(\frac{\partial \bar{L}}{\partial \ddot{\varphi}} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \bar{L}}{\partial \dddot{\varphi}}\right)\ddot{\varphi} \\ & + \left(\frac{\partial \bar{L}}{\partial \dot{\varphi}} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \bar{L}}{\partial \ddot{\varphi}} + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial \bar{L}}{\partial \dddot{\varphi}}\right)\dot{\varphi} - \bar{L}, \end{aligned} \tag{2.69}$$

by using the all-orders two-oscillator drift kinetic equations of motion (2.20)–(2.22) to express the bare Hamiltonian in terms of the true phase-space variables $(Q, P, \varphi, \dot{\varphi})$. A straightforward calculation, similar to the calculation of $\tilde{\vartheta}^\infty$, shows that the dressed Hamiltonian is given by

$$\tilde{H}^\infty(Q, P, \varphi, \dot{\varphi}) = \tfrac{1}{2}\dot{\varphi}^2 + \tfrac{1}{2}P^2 + \tfrac{1}{2}\omega_2^2\varphi^2 + \tfrac{1}{2}\hat{q}_-^2(Q, \varphi)(\omega_1^2 + 2g\varphi). \tag{2.70}$$

By applying fixed-endpoint variations to the phase-space action, it is possible, yet tedious, to verify directly that all-orders two-oscillator drift kinetics admits this alternative variational principle. Indeed, for a general dressed Hamiltonian the Euler–Lagrange equations give

$$\frac{\mathrm{d}}{\mathrm{d}t}Q = \sqrt{D}\frac{\partial \tilde{H}^\infty}{\partial P} - \epsilon\omega_2^2\frac{\partial \tilde{H}^\infty}{\partial \dot{\varphi}} \tag{2.71}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}P = -\sqrt{D}\frac{\partial \tilde{H}^\infty}{\partial Q} \tag{2.72}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi = \frac{\partial \tilde{H}^\infty}{\partial \dot{\varphi}} \tag{2.73}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\dot{\varphi} = -\frac{\partial \tilde{H}^\infty}{\partial \varphi} + \epsilon\omega_2^2\frac{\partial \tilde{H}^\infty}{\partial Q}. \tag{2.74}$$

When the physical dressed Hamiltonian given in (2.70) is substituted into this last equation set, the dynamical equations defining all-orders two-oscillator drift kinetics (on the bottom sheet) are recovered.

With this phase-space action principle for all-orders two-oscillator drift kinetics in hand, I will now use it to construct an alternative truncated theory that I will refer to as renormalized two-oscillator drift kinetics (I will explain the use of the term 'renormalized' shortly). Renormalized two-oscillator drift kinetics is defined simply by replacing $\tilde{\vartheta}^\infty$ and $\tilde{H}^\infty$ in (2.68) with the corresponding first-order Taylor expansions in ϵ, $\tilde{\vartheta}$ and $\tilde{H}$, where

$$\tilde{\vartheta} = P(1 + \epsilon 2gQ)\,\mathrm{d}Q + (\dot{\varphi} + \epsilon P\omega_2^2)\,\mathrm{d}\varphi \tag{2.75}$$

$$\begin{aligned} \tilde{H} = {} & \tfrac{1}{2}\dot{\varphi}^2 + \tfrac{1}{2}P^2 + \tfrac{1}{2}\omega_2^2\varphi^2 + \tfrac{1}{2}(\omega_1^2 + 2g\varphi)Q^2 \\ & + \epsilon(\omega_1^2\omega_2^2 Q\varphi + 2g\omega_2^2 Q\varphi^2 + [2g^2\varphi + g\omega_1^2]Q^3). \end{aligned} \tag{2.76}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

Thus the phase-space action for renormalized two-oscillator drift kinetics is given by

$$S(\gamma) = \int_{\gamma} \tilde{\vartheta} - \int_{t_1}^{t_2} \tilde{H}(\gamma(t))\,\mathrm{d}t, \tag{2.77}$$

and the associated equations of motion are

$$\frac{\mathrm{d}}{\mathrm{d}t}Q = \frac{1}{1+2\epsilon gQ}\left(\frac{\partial \tilde{H}}{\partial P} - \epsilon\omega_2^2\frac{\partial \tilde{H}}{\partial \dot{\varphi}}\right) \tag{2.78}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}P = -\frac{1}{1+2\epsilon gQ}\frac{\partial \tilde{H}}{\partial Q} \tag{2.79}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi = \frac{\partial \tilde{H}}{\partial \dot{\varphi}} \tag{2.80}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\dot{\varphi} = -\frac{\partial \tilde{H}}{\partial \varphi} + \frac{\epsilon\omega_2^2}{1+2\epsilon gQ}\frac{\partial \tilde{H}}{\partial Q}. \tag{2.81}$$

It is simple to check that the dynamical vector field for renormalized two-oscillator drift kinetics agrees with the 'bottom sheet' dynamical vector field for all-orders two-oscillator drift kinetics with $O(\epsilon)$ accuracy. Moreover, renormalized two-oscillator drift kinetics is superior to two-oscillator drift kinetics because it is defined on the correct phase space. This last point is illustrated clearly by considering the linearization of renormalized two-oscillator drift kinetics about the equilibrium $(Q, P, \varphi, \dot{\varphi}) = 0$. The linearized equations are given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta Q = \delta P - \epsilon\omega_2^2\delta\dot{\varphi} \tag{2.82}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta P = -\omega_1^2(\delta Q + \epsilon\omega_2^2\delta\varphi) \tag{2.83}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta\varphi = \delta\dot{\varphi} \tag{2.84}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta\dot{\varphi} = -\omega_2^2(1-\epsilon^2\omega_1^2\omega_2^2)\delta\varphi, \tag{2.85}$$

which can be written as $\dot{z} = Az$, where $z = (\delta Q, \delta P, \delta\varphi, \delta\dot{\varphi})$ and A is a 4×4 matrix. The spectrum of A is given by $\{\mathrm{i}\omega_1, -\mathrm{i}\omega_1, \mathrm{i}\omega_2\sqrt{1-\epsilon^2\omega_1^2\omega_2^2}, -\mathrm{i}\omega_2\sqrt{1-\epsilon^2\omega_1^2\omega_2^2}\}$, which implies that all normal modes of linearized renormalized two-oscillator drift kinetics are oscillatory, just as in the all-orders theory. Moreover, the normal mode frequencies of the renormalized theory are close to those of the all-orders theory.

How did this modified Lagrangian procedure succeed in preserving the correct phase space? The answer is that by implementing approximations at the level of the system phase-space Lagrangian it is impossible for the phase space of all-orders two-oscillator drift kinetics to be modified (modulo non-perturbative modifications of the phase-space sheet structure). As long as the 1-form appearing in the phase-space action has a non-degenerate exterior derivative, the Euler–Lagrange equations that follow from the phase-space action principle are guaranteed to define a first-order ordinary differential equation (ODE) on the space where the 1-form is defined. This non-degeneracy condition is morally guaranteed by the fact that the transformation between the particle coordinates and the guiding centre coordinates is near-identity.

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

In what sense is renormalized two-oscillator drift kinetics a 'renormalized' theory? While I do not believe the renormalization group is playing any role here, the procedure described in this subsection bears a strong resemblance to the procedure used to extract sensible results from perturbative quantum field theories. I began with a system with a nonsensical phase space, namely two-oscillator drift kinetics. In a loose sense, this system contained an unphysical infinity. To eliminate the absurd features of this model, I regularized it by passing to all-orders two-oscillator drift kinetics. This step can be thought of as inserting a 'cutoff'; I inserted higher-order ϵ-dependent terms into the two-oscillator drift kinetic Lagrangian that guaranteed the dimension of phase space was sensible. Finally, I devised a method to vary the cutoff (i.e. change which higher-order terms were in the theory) that would not alter the value of the dimension of phase space. This step is analogous to tailoring the values of the coupling constants in a perturbative quantum field theory (Wentzel 1949) in order to render the propagator cutoff-independent.

3. Conventional Lagrangian electrostatic drift kinetics and its flaws

In passing from two-oscillator kinetics to two-oscillator drift kinetics, the size of phase space was increased for unphysical reasons. In this section I will argue that the same phenomenon occurs in a variant of first-order high-flow electrostatic drift kinetics that I will refer to as MSST drift kinetics. MSST drift kinetics is defined by the system Lagrangian

$$L(F, V, \varphi, \dot{\varphi}) = \sum_s \int \ell_s(z, V_s(z); \varphi, \dot{\varphi})\, F_s \, \mathrm{d}\Omega_s + \frac{\epsilon_o}{2} \int |\nabla \varphi|^2 \, \mathrm{d}^3 \boldsymbol{X}, \tag{3.1}$$

where the guiding centre phase-space Lagrangian ℓ_s is the one given in equation (45) of Miyato *et al.* (2009),

$$\ell_s(z, \dot{z}; \varphi, \dot{\varphi}) = \left(q_s \boldsymbol{A} + m_s v_\parallel \hat{b} - \frac{m_s}{q_s} \mu \boldsymbol{W} \right) \cdot \dot{\boldsymbol{X}} + \frac{q_s}{m_s} \mu \, \dot{\zeta} - h_s(z; \varphi, \boldsymbol{E}, \dot{\boldsymbol{E}}), \tag{3.2}$$

with

$$h_s(z; \varphi, \boldsymbol{E}, \dot{\boldsymbol{E}}) = q_s \varphi + K_s(\boldsymbol{E}, \dot{\boldsymbol{E}}), \tag{3.3}$$

and $K_s(\boldsymbol{E}, \dot{\boldsymbol{E}}) = K_{os}(\boldsymbol{E}) + K_{1s}(\boldsymbol{E}, \dot{\boldsymbol{E}})$ is given by

$$K_{os}(\boldsymbol{E}) = \tfrac{1}{2} m_s v_\parallel^2 + \mu B - \tfrac{1}{2} m_s u_E^2 \tag{3.4}$$

$$\begin{aligned} K_{1s}(\boldsymbol{E}, \dot{\boldsymbol{E}}) = {} & \frac{1}{2} \frac{m_s}{q_s} \left(\mu + \frac{m_s u_E^2}{2B} \right) \hat{b} \cdot \nabla \times \boldsymbol{u}_E \\ & - \frac{m_s v_\parallel}{2\omega_{cs}} \boldsymbol{u}_E \cdot \nabla \times \boldsymbol{u}_E + \frac{7}{6} \frac{m_s}{q_s} \mu (\hat{b} \times \boldsymbol{u}_E) \cdot \nabla \ln B \\ & + \frac{m_s u_E^2 v_\parallel}{\omega_{cs}} \tau + \frac{m_s v_\parallel^2}{\omega_{cs}} \hat{b} \times \boldsymbol{u}_E \cdot \kappa - \frac{q_s}{2\omega_{cs}^2} \boldsymbol{u}_E \cdot \dot{\boldsymbol{E}}. \end{aligned} \tag{3.5}$$

Here F is the (multi-species) distribution function; V is the Eulerian phase-space fluid velocity; φ is the dynamical electrostatic potential; $z = (\boldsymbol{X}, v_\parallel, \mu, \zeta)$ is a point in the guiding centre phase space consisting of the guiding centre position, parallel

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

velocity, magnetic moment and gyrophase; $\boldsymbol{A}$ is the background vector potential; $\hat{b}$ is the unit vector along the background magnetic field; $\tau = \hat{b} \cdot \boldsymbol{\nabla} \times \hat{b}$ and $\kappa = \hat{b} \cdot \boldsymbol{\nabla} \hat{b}$; $\boldsymbol{W} = (\boldsymbol{\nabla} \hat{e}_1) \cdot \hat{e}_2 + \tau \hat{b}/2$, where $\hat{e}_1, \hat{e}_2$ are unit vectors perpendicular to the background magnetic field; $\boldsymbol{E} = -\boldsymbol{\nabla}\varphi$ is the electric field; $\boldsymbol{u}_E = \boldsymbol{E} \times \hat{b}/B$ is the $E \times B$ velocity; the Liouville volume element $\mathrm{d}\Omega_s = \mathcal{J}_s \, \mathrm{d}^3\boldsymbol{X} \, \mathrm{d}v_\parallel \, \mathrm{d}\mu \, \mathrm{d}\zeta$; and $\mathcal{J}_s = B^*_{\parallel s}/m_s$, where $\boldsymbol{B}^*_s = \boldsymbol{\nabla} \times (\boldsymbol{A} + m_s v_\parallel \hat{b}/q_s - m_s \mu \boldsymbol{W}/q_s^2)$. The authors of (Miyato *et al.* 2009) drop a number of terms from K_{1s} by invoking some often-relevant subsidiary orderings (relative to the usual $\epsilon = \rho/L$ ordering). Here, I will not invoke these subsidiary orderings, and will therefore keep all of the terms in K_{1s}.

The equations of motion corresponding to this system Lagrangian may be obtained by varying φ, F and V in the action integral $S = \int_{t_1}^{t_2} L \, \mathrm{d}t$ subject to the constraints

$$\delta V_s = \dot{\xi}_s + L_{V_s} \xi_s \tag{3.6}$$

$$\delta F_s = -\frac{1}{\mathcal{J}_s} \mathrm{div}(\mathcal{J}_s \xi_s F_s), \tag{3.7}$$

where ξ is an arbitrary time-dependent (multi-species) vector field on the guiding centre phase space and L_{V_s} denotes the Lie derivative along the vector field V_s. A thorough understanding of the rationale behind using these constrained variations can be obtained from a reading of the original article on Euler–Poincaré variational principles, namely Holm, Marsden & Ratiu (1998), and the first articles to present Euler–Poincaré formulations of the Vlasov–Maxwell system (Cendra *et al.* 1998) and an electrostatic gyrokinetic system (Squire *et al.* 2013). Varying φ gives the Poisson equation,

$$\boldsymbol{\nabla} \cdot \left(\epsilon_o \boldsymbol{E} - \frac{\delta \mathcal{K}}{\delta \boldsymbol{E}} + \partial_t \frac{\delta \mathcal{K}}{\delta \dot{\boldsymbol{E}}} \right) = \sum_s q_s N_s, \tag{3.8}$$

where $\mathcal{K} = \mathcal{K}(F, \boldsymbol{E}, \dot{\boldsymbol{E}})$ is given by

$$\mathcal{K}(F, \boldsymbol{E}, \dot{\boldsymbol{E}}) = \sum_s \int K_s(\boldsymbol{E}, \dot{\boldsymbol{E}}) F_s \, \mathrm{d}\Omega_s, \tag{3.9}$$

$$N_s = \int F_s \, \mathcal{J}_s \, \mathrm{d}v_\parallel \, \mathrm{d}\mu \, \mathrm{d}\zeta, \tag{3.10}$$

and $\omega_{cs} = q_s B/m_s$. Varying V and F gives the guiding centre equations of motion

$$V_s^i = \{z^i, h_s\}_s^{gc}, \tag{3.11}$$

where $\{\cdot, \cdot\}^{gc}$ denotes the guiding centre Poisson bracket associated with the guiding centre phase-space Lagrangian ℓ_s. Explicitly,

$$\dot{\boldsymbol{X}}_s = \frac{\boldsymbol{B}^*_s}{m_s B^*_{\parallel s}} \frac{\partial h_s}{\partial v_\parallel} + \frac{\hat{b} \times \boldsymbol{\nabla} h_s}{q_s B^*_{\parallel s}} \tag{3.12}$$

$$\dot{v}_{\parallel s} = -\frac{\boldsymbol{B}^*_s \cdot \boldsymbol{\nabla} h_s}{m_s B^*_{\parallel s}}. \tag{3.13}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

The distribution function F therefore satisfies

$$\frac{\partial F_s}{\partial t} + \frac{1}{\mathcal{J}_s}\nabla\cdot(\dot{\boldsymbol{X}}_s F_s \mathcal{J}_s) + \frac{1}{\mathcal{J}_s}\frac{\partial}{\partial v_\parallel}(\dot{v}_{\parallel s} F_s \mathcal{J}_s) = 0. \tag{3.14}$$

The dynamical system specified by (3.8) and (3.14) is meant to approximate the Vlasov–Poisson system in a strong background magnetic field when the time and length scales of the electrostatic potential are long compared with the gyrofrequency and gyroradius, respectively. In particular, because the electrostatic potential in the Vlasov–Poisson system is a functional of the distribution function, (3.8) and (3.14) are supposed to approximate a system of equations whose phase space is the space of F's. Moreover, I will argue in the next section that when high-flow electrostatic gyrokinetics is carried to all orders in magnetic non-uniformity, the phase space is indeed the space of F's. It is therefore reasonable to hope that (3.8) and (3.14) are equivalent to a first-order ODE on the space of F's.

Unfortunately this hope is ill-founded. I will explain why in the context of linearized MSST drift kinetics. For simplicity's sake, let the background magnetic field be uniform and z-directed. The equilibrium I will linearize about satisfies $\varphi = \dot{\varphi} = 0$ and

$$F_{os} = \left(\frac{m_s}{2\pi T}\right)^{3/2} N_{os}(\boldsymbol{X})\exp(-(\mu B + m_s v_\parallel^2/2)/T), \tag{3.15}$$

where T is the equilibrium temperature and N_{os} is the equilibrium number density. This type of equilibrium will exist if and only if

$$-\nabla\cdot\nabla_\perp\left(\sum_s q_s\rho_s^2 N_{os}\right) = \sum_s q_s N_{os} \tag{3.16}$$

$$\hat{b}\cdot\nabla N_{os} = 0, \tag{3.17}$$

where $\rho_s^2 \equiv (1/2)((T/m_s)/\omega_{cs}^2)$ is a typical squared gyroradius of a species-s particle. Linearization of MSST drift kinetics about this type of equilibrium leads to the dynamical equations

$$-\epsilon_o\Delta_\perp\delta\varphi = \sum_s q_s\delta N_s + \nabla_\perp^2\left(\sum_s\frac{\delta M_s}{2\omega_{cs}}\right) + \epsilon_o\nabla\tau_o\times\hat{b}\cdot\nabla\delta\dot{\varphi} \tag{3.18}$$

$$\delta\dot{F}_s + v_\parallel\hat{b}\cdot\nabla\delta F_s + \frac{\hat{b}}{q_s B}\times\nabla\left(q_s\delta\varphi + \frac{\mu\nabla_\perp^2\delta\varphi}{2\omega_{cs}}\right)\cdot\nabla F_{so} - \frac{\hat{b}}{m_s}\cdot\nabla\left(q_s\delta\varphi + \frac{\mu\nabla_\perp^2\delta\varphi}{2\omega_{cs}}\right)\frac{\partial F_{os}}{\partial v_\parallel} = 0, \tag{3.19}$$

where the elliptic linear operators $\Delta_\perp$ and $\nabla_\perp^2$ are given by

$$\Delta_\perp u = \Delta u + \nabla\cdot\left(\sum_s\frac{c^2\mu_o m_s N_{os}}{B^2}\nabla_\perp u\right) \tag{3.20}$$

$$\nabla_\perp^2 u = \nabla\cdot\nabla_\perp u, \tag{3.21}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

and the quantities δM_s, τ_o are the perturbed perpendicular pressure,

$$\delta M_s = \int \mu \delta F_s \frac{B}{m_s} \, \mathrm{d}v_\parallel \, \mathrm{d}\mu \, \mathrm{d}\zeta \,, \tag{3.22}$$

and

$$\tau_o = \sum_s \frac{q_s N_{os}}{\epsilon_o \omega_{cs}^2 B}. \tag{3.23}$$

Instead of analysing these equations in their full generality, I will restrict attention to proton–electron plasmas with equilibrium density given by

$$N_{os} = n\left(1 + \frac{x}{L}\right), \tag{3.24}$$

that are uniform along the z-axis. In this special case, the linearized MSST equations can be reduced to a system of equations involving only the perturbed electrostatic potential,

$$-\epsilon_o \Delta_\perp \delta\dot{\varphi} - (\gamma \nabla_\perp^4 + \alpha \nabla_\perp^2)\partial_y \delta\varphi + \beta \partial_y \delta\ddot{\varphi} = 0, \tag{3.25}$$

where

$$\alpha = \sum_s \frac{nT}{B^2 \omega_{cs} L} \tag{3.26}$$

$$\beta = \sum_s \frac{m_s n}{B^2 \omega_{cs} L} \tag{3.27}$$

$$\gamma = \sum_s \frac{nT^2}{2B^2 \omega_{cs}^3 L m_s}, \tag{3.28}$$

$$\Delta_\perp \delta\dot{\varphi} = \frac{c^2}{v_A^2} \nabla_\perp^2 \delta\dot{\varphi} + \frac{c^2}{v_{Ao}^2 L} \partial_x \delta\dot{\varphi}, \tag{3.29}$$

and

$$\frac{c^2}{v_{Ao}^2} = \sum_s \frac{c^2 \mu_o m_s n}{B^2} \tag{3.30}$$

$$\frac{c^2}{v_A^2} = 1 + \frac{c^2}{v_{Ao}^2}\left(1 + \frac{x}{L}\right). \tag{3.31}$$

A thorough analysis of (3.25), which is linear with non-constant coefficients, requires specifying boundary conditions on $\delta\varphi$. However, in order to reveal the nature of modes that vary rapidly in space and time, it is enough to find the dispersion relation in the Wentzel–Kramers–Brillouin (WKB) approximation. Set $\delta\varphi = a \exp(iS)$ where a is a slowly varying amplitude and S is a rapidly varying phase function. Define $\omega = \partial_t S$ and $\boldsymbol{k} = \boldsymbol{\nabla} S$. Upon substituting this ansatz into (3.25) and then dropping all terms involving gradients of the amplitude a, the following dispersion relation emerges

$$-\beta k_y \omega^2 + \epsilon_o \left(\frac{c^2}{v_A^2}(k_x^2 + k_y^2) + \frac{c^2}{v_{Ao}^2}\frac{k_x}{iL}\right)\omega + (\alpha(k_x^2 + k_y^2) + \gamma(k_x^2 + k_y^2)^2)k_y = 0. \tag{3.32}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

Upon noting that $\alpha, \beta = O(\epsilon)$ and that $\gamma = O(\epsilon^3)$, it becomes clear that when $k_x \ll k_y$ the two branches of this dispersion relation are given to leading order in ϵ as

$$\omega_d \approx -\frac{v_* k_y}{1 + \frac{x}{L} + \frac{v_{Ao}^2}{c^2}} \tag{3.33}$$

$$\omega_\times \approx k_y \frac{T/m_i}{v_*}\left(1 + \frac{x}{L} + \frac{v_{Ao}^2}{c^2}\right), \tag{3.34}$$

where $v_* = T/(q_i BL) = O(\epsilon)$ is the so-called diamagnetic drift velocity. The first branch, ω_d, reproduces the well-known drift wave dispersion relation. The second branch, $\omega_\times$, is unphysical because its frequency increases as $\epsilon \to 0$. This implies that the wave violates the drift kinetic approximation, namely $\omega \ll \omega_c$, where ω is any characteristic frequency of interest. The unphysical branch is ushered into the theory by the appearance of the second time derivative of $\delta\varphi$ in (3.25), which itself appears as a consequence of the first time derivative of $\delta\varphi$ appearing in the linearized drift kinetic Poisson equation, (3.18). In short, the presence of the unphysical mode is a direct consequence of the impossibility of expressing MSST drift kinetics as a first-order ODE in F-space. In fact we have seen that MSST drift kinetics reduces to a first-order ODE in (φ, F)-space when linearized about the class of slab equilibria just considered, which is larger than F-space in the sense that F-space is a proper subset of (φ, F)-space. In the following section I will construct an alternative variational drift kinetic theory that is just as accurate (in terms of ϵ) as MSST drift kinetics, but that does not suffer from an unphysically large phase space and its associated unphysical rapidly varying modes.

4. Renormalized MSST drift kinetics

4.1. *The all-orders phase-space Lagrangian*

The all-orders version of MSST drift kinetics is derived as follows. The starting point is the magnetized Vlasov–Poisson system expressed in a special coordinate system on single-particle phase space,

$$\partial_t f_s + \nabla \cdot ([\boldsymbol{u}_E + \boldsymbol{w}] f_s) + \nabla_{\boldsymbol{w}} \cdot (\dot{\boldsymbol{w}}_s f_s) = 0 \tag{4.1}$$

$$\epsilon_o \Delta\varphi = -\rho(f) \tag{4.2}$$

$$\dot{\boldsymbol{w}}_s = -\partial_t \boldsymbol{u}_E - \boldsymbol{u}_E \cdot \nabla \boldsymbol{u}_E - \boldsymbol{w} \cdot \nabla \boldsymbol{u}_E + \frac{q_s}{m_s} \boldsymbol{w} \times \boldsymbol{B}, \tag{4.3}$$

where $\rho(f) = \sum_s q_s \int f_s \, \mathrm{d}^3 \boldsymbol{w}$, $\boldsymbol{B}$ is a static nowhere vanishing magnetic field, the velocity-like variable $\boldsymbol{w}$ is related to the single-particle velocity $\boldsymbol{v}$ by $\boldsymbol{v} = \boldsymbol{u}_E + \boldsymbol{w}$, and f is the distribution function on $(\boldsymbol{x}, \boldsymbol{w})$-space. For simplicity, I will assume the spatial domain is all of $\mathbb{R}^3$ and that we seek square integrable electrostatic potentials φ with square integrable gradients. In this space of admissible φ, the Laplacian Δ is invertible, and I will denote its inverse $G = \Delta^{-1}$. Using G, the magnetized Vlasov–Poisson system can be expressed as the ODE on f-space,

$$\partial_t f_s = -\nabla \cdot (\dot{\boldsymbol{x}}(f) f_s) - \nabla_{\boldsymbol{w}} \cdot (\dot{\boldsymbol{w}}_s(f) f_s) \tag{4.4}$$

$$\equiv \dot{f}_s(f), \tag{4.5}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

where

$$\dot{\boldsymbol{x}}(f) = \boldsymbol{u}_E(f) + \boldsymbol{w} \tag{4.6}$$

$$\dot{\boldsymbol{w}}_s(f) = -\dot{\boldsymbol{u}}_E(f) - \boldsymbol{u}_E(f)\cdot\nabla\boldsymbol{u}_E(f) - \boldsymbol{w}\cdot\nabla\boldsymbol{u}_E(f) + \frac{q_s}{m_s}\boldsymbol{w}\times\boldsymbol{B}, \tag{4.7}$$

and the vectors $\boldsymbol{u}_E(f) = \boldsymbol{\mathcal{E}}(f)\times b/B$ and $\dot{\boldsymbol{u}}_E(f) = \boldsymbol{\mathcal{E}}^{(1)}(f)\times b/B$ are determined by the operators

$$\boldsymbol{\mathcal{E}}(f) = \frac{1}{\epsilon_o}\nabla G[\rho(f)] \tag{4.8}$$

$$\boldsymbol{\mathcal{E}}^{(1)}(f) = -\frac{1}{\epsilon_o}\nabla G[\nabla\cdot(\boldsymbol{u}_E(f)\,\rho(f) + \boldsymbol{J}(f))], \tag{4.9}$$

with $\boldsymbol{J}(f) = \sum_s q_s\int \boldsymbol{w} f_s\,\mathrm{d}^3\boldsymbol{w}$. Note that the magnetized Vlasov–Poisson equations imply $\partial_t\rho(f) = -\nabla\cdot(\boldsymbol{u}_E(f)\rho(f) + \boldsymbol{J}(f))$.

Next, a near-identity field-dependent phase-space coordinate transformation, $\Phi_s(j\boldsymbol{E})\!:\!P\to P$, is identified that decouples the rapid gyromotion from the slower drift motion. Here $j\boldsymbol{E} = (\boldsymbol{E}, \partial_t\boldsymbol{E}, \partial_t^2\boldsymbol{E}, \ldots)$ is an infinite sequence of vector fields, and P is the six-dimensional position-velocity phase space. There are many possible ways to choose this coordinate transformation, and MSST drift kinetics is defined by the particular choice described in Miyato *et al.* (2009).

Using this coordinate transformation, the guiding centre distribution function F is defined as

$$F_s = \Phi_s(j\boldsymbol{E})_* f_s, \tag{4.10}$$

where f_s is the particle distribution function, and the subscripted $*$ denotes the pushforward. Assuming that $\boldsymbol{E}$ and f_s satisfy the magnetized Vlasov–Poisson equations, the right-hand-side of (4.10) can be expressed entirely in terms of f because $j\boldsymbol{E} = (j\boldsymbol{\mathcal{E}})(f) = (\boldsymbol{\mathcal{E}}(f), \boldsymbol{\mathcal{E}}^{(1)}(f), \boldsymbol{\mathcal{E}}^{(2)}(f), \ldots)$, where

$$\boldsymbol{\mathcal{E}}(f) = \frac{1}{\epsilon_o}\nabla G[\rho(f)] \tag{4.11}$$

$$\boldsymbol{\mathcal{E}}^{(1)}(f) = -\frac{1}{\epsilon_o}\nabla G[\nabla\cdot(\boldsymbol{u}_E(f)\,\rho(f) + \boldsymbol{J}(f))] \tag{4.12}$$

$$\boldsymbol{\mathcal{E}}^{(2)}(f) = \ldots, \tag{4.13}$$

and the operators $\mathcal{E}^{(n)}$ for $n>1$ are determined by expressing the nth time derivative of $\boldsymbol{E}$ in terms of f using the magnetized Vlasov–Poisson equations. To be precise,

$$F_s = \Phi_s(j\boldsymbol{\mathcal{E}}(f))_* f_s \tag{4.14}$$

$$\equiv \mathcal{F}_s^{-1}(f). \tag{4.15}$$

Note that $\mathcal{F}^{-1}$ maps multi-species functions on P into multi-species functions on P; formally $\mathcal{F}^{-1}\!:\oplus_s C^\infty(P)\to\oplus_s C^\infty(P)$, where $\oplus_s$ denotes the species-sum of vector spaces.

Because Φ is near-identity, $\mathcal{F}_s^{-1}(f)\approx f_s$. This implies that $\mathcal{F}^{-1}$ has a unique perturbatively defined inverse $\mathcal{F}$ that satisfies

$$F_s = \Phi_s(j\boldsymbol{\mathcal{E}}(\mathcal{F}(F)))_*\mathcal{F}_s(F). \tag{4.16}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

I will explain the procedure for obtaining an asymptotic expansion of $\mathcal{F}$ in the next subsection. Therefore,

$$\partial_t \mathcal{F}(F) = \partial_t f \tag{4.17}$$

$$\Rightarrow D\mathcal{F}[\partial_t F] = \dot{f}(f) \tag{4.18}$$

$$\Rightarrow \partial_t F = D\mathcal{F}^{-1}[\dot{f}(\mathcal{F}(F))]. \tag{4.19}$$

Here $D\mathcal{F}$ denotes the Fréchet derivative of $\mathcal{F}$. Assuming that Φ is chosen according to the prescription given in Miyato *et al.* (2009), (4.19) is the all-orders version of MSST drift kinetics. Because this system is manifestly an ODE on F-space, the phase space for all-orders MSST drift kinetics is F-space, which proves my earlier claim.

A variational principle for all-orders MSST drift kinetics can be derived by transforming the following variational principle for the magnetized Vlasov–Poisson system into guiding centre coordinates. The Vlasov–Poisson Lagrangian is given by

$$\mathfrak{L}(f, U, \varphi, \boldsymbol{E}) = \sum_s \int l_s(z, U_s(z), \varphi, \boldsymbol{E}) f_s \, \mathrm{d}\omega + \frac{\epsilon_o}{2} \int |\nabla\varphi|^2 \, \mathrm{d}^3\boldsymbol{x}, \tag{4.20}$$

where the single-particle phase-space Lagrangian is

$$l_s(z, \dot{z}, \varphi, \boldsymbol{E}) = (m_s \boldsymbol{w} + m_s \boldsymbol{u}_E + q_s \boldsymbol{A}) \cdot \dot{\boldsymbol{x}} - H_s(z, \varphi, \boldsymbol{E}), \tag{4.21}$$

the single-particle Hamiltonian is

$$H_s(z, \varphi, \boldsymbol{E}) = \tfrac{1}{2} m_s (\boldsymbol{w} + \boldsymbol{u}_E)^2 + q_s \varphi, \tag{4.22}$$

the volume element $\mathrm{d}\omega = \mathrm{d}^3\boldsymbol{x}\, \mathrm{d}^3\boldsymbol{w}$, and U is the Eulerian phase-space fluid velocity. The equations of motion corresponding to this system Lagrangian may be obtained by varying φ, f and U in the action integral $S = \int_{t_1}^{t_2} L \, \mathrm{d}t$ subject to the constraints

$$\delta U_s = \dot{\xi}_s + L_{U_s} \xi_s \tag{4.23}$$

$$\delta f_s = -\mathrm{div}(\xi_s f_s), \tag{4.24}$$

where ξ is an arbitrary time-dependent vector field on the (multi-species) guiding centre phase space and L_{U_s} denotes the Lie derivative along the vector field U_s.

Using the MSST prescription for choosing Φ, the transformed system Lagrangian is given by

$$L^\infty(F, V, \varphi) = \sum_s \int \ell_s^\infty(z, V_s(z), \varphi, j\boldsymbol{E}) \, F_s \, \mathrm{d}\Omega_s + \frac{\epsilon_o}{2} \int |\nabla\varphi|^2 \, \mathrm{d}^3\boldsymbol{x}, \tag{4.25}$$

where the all-orders guiding centre phase-space Lagrangian is given by

$$\ell_s^\infty(z, \dot{z}, \varphi, j\boldsymbol{E}) = \left(q_s \boldsymbol{A} + m_s v_\parallel \hat{b} - \frac{m_s}{q_s} \mu \boldsymbol{W} \right) \cdot \dot{\boldsymbol{X}} + \frac{q_s}{m_s} \mu \, \dot{\zeta} - h_s^\infty(z, \varphi, j\boldsymbol{E}), \tag{4.26}$$

and the all-orders guiding centre Hamiltonian is of the form

$$h_s^\infty(z, \varphi, j\boldsymbol{E}) = K_s^\infty(j\boldsymbol{E}) + q_s \varphi, \tag{4.27}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

where $K_s^\infty = K_{os} + K_{1s} + K_{2s} + \cdots$ and K_{os}, K_{1s} are given in (3.4) and (3.5). Note that the MSST method for finding the guiding centre transformation Φ produces a guiding centre phase-space Lagrangian with a simple time-independent 'symplectic part'. The symplectic part of ℓ_s^∞, ℓ_s^{symp}, is

$$\begin{aligned}\ell_s^{symp}(z, \dot{z}) &= \ell_s^\infty + h_s^\infty \\ &= \left(q_s \boldsymbol{A} + m_s v_\parallel \hat{b} - \frac{m_s}{q_s}\mu \boldsymbol{W}\right) \cdot \dot{\boldsymbol{X}} + \frac{q_s}{m_s}\mu\dot{\zeta},\end{aligned} \tag{4.28}$$

and achieving this time-independence was one of the principal goals of Miyato *et al.* (2009). The variational principle associated with L^∞ is very similar to the one introduced earlier for first-order MSST drift kinetics. The action $S^\infty = \int_{t_1}^{t_2} L^\infty \,\mathrm{d}t$ is varied subject to the constraints given by

$$\delta V_s = \dot{\xi}_s + L_{V_s}\xi_s \tag{4.29}$$

$$\delta F_s = -\frac{1}{\mathcal{J}_s}\mathrm{div}(\mathcal{J}_s \xi_s F_s). \tag{4.30}$$

Using this variational principle for all-orders MSST drift kinetics, the all-orders system phase-space Lagrangian can finally be derived using the method described in § 2. The dressed Lagrange 1-form and the dressed Hamiltonian must each be found.

The first step toward finding the dressed Lagrange 1-form $\tilde{\vartheta}^\infty$ is to identify the bare Lagrange 1-form ϑ^∞. This is done by examining the endpoint contributions to free-endpoint variations of S^∞. A straightforward calculation shows that the first free-endpoint variation of S^∞ about a solution of all-orders MSST drift kinetics is given by

$$\delta S^\infty = \vartheta^\infty[\xi, j\delta \boldsymbol{E}]|_{t_1}^{t_2}, \tag{4.31}$$

where the bare Lagrange 1-form ϑ^∞ is given by

$$\begin{aligned}\vartheta^\infty[\xi, \delta j\boldsymbol{E}] &= v\sum_s \int \ell_s^{symp}(z, \xi_s(z))\, F_s \,\mathrm{d}\Omega_s \\ &\quad - \sum_{k=1}^{\infty}\sum_{j=1}^{k}\int (\delta \partial_t^{k-j}\boldsymbol{E}) \cdot \left((-\partial_t)^{j-1}\frac{\delta \mathcal{K}^\infty}{\delta \partial_t^k \boldsymbol{E}}\right) \mathrm{d}^3\boldsymbol{X},\end{aligned} \tag{4.32}$$

with $\mathcal{K}^\infty(F, j\boldsymbol{E}) = \sum_s \int K_s^\infty(j\boldsymbol{E})\, F_s \,\mathrm{d}\Omega_s$. In deriving this expression, the all-orders Euler–Lagrange equations,

$$V_s^{\infty i} = \{z^i, h_s^\infty\}_s^{gc} \tag{4.33}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{D} = \rho(F) \tag{4.34}$$

$$\boldsymbol{D} = \epsilon_o \boldsymbol{E} - \frac{\delta \mathcal{K}^\infty}{\delta \boldsymbol{E}} - \sum_{k=1}^{\infty}(-\partial_t)^k \frac{\delta \mathcal{K}^\infty}{\delta \partial_t^k \boldsymbol{E}}, \tag{4.35}$$

must be used to eliminate all contributions to δS^∞ except for those that come from the endpoints. The dressed Lagrange 1-form $\tilde{\vartheta}$ can now be obtained by expressing ϑ

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

in terms of F and ξ using the all-orders MSST drift kinetic system. All of the time derivatives of $\boldsymbol{E}$ must be expressed in terms of F using the relationship

$$j\boldsymbol{E} = j\boldsymbol{\mathcal{E}}(\mathcal{F}(F)) \equiv j\tilde{\boldsymbol{\mathcal{E}}}(F). \tag{4.36}$$

Similarly, the variations of the time derivatives of $\boldsymbol{E}$ must be expressed in terms of ξ using

$$\partial_t^{k-j}\boldsymbol{E} = \boldsymbol{\mathcal{E}}^{(k-j)}(\mathcal{F}(F)) \equiv \tilde{\boldsymbol{\mathcal{E}}}^{(k-j)}(F), \tag{4.37}$$

which gives

$$\begin{aligned}\delta\partial_t^{k-j}\boldsymbol{E} &= -D\tilde{\boldsymbol{\mathcal{E}}}^{(k-j)}[\mathcal{J}^{-1}\mathrm{div}(\mathcal{J}F\xi)] \\ &= -\sum_s D_s\tilde{\boldsymbol{\mathcal{E}}}^{(k-j)}[\mathcal{J}_s^{-1}\mathrm{div}(\mathcal{J}_sF_s\xi_s)].\end{aligned} \tag{4.38}$$

Here D_s denotes a partial Fréchet derivative with respect to the species-s distribution function. When this dressing procedure is complete, the all-orders dressed Lagrange 1-form can be expressed as

$$\tilde{\vartheta}^\infty[\xi] = \sum_s \int \tilde{\ell}_s^{symp,\infty}(z, \xi_s(z); F)\, F_s\, \mathrm{d}\Omega_s, \tag{4.39}$$

where

$$\tilde{\ell}_s^{symp,\infty}(z, \dot{z}; F) = \ell_s^{symp}(z, \dot{z}) + \dot{z}^i\partial_{z^i}\alpha_s^\infty(F), \tag{4.40}$$

and the function $\alpha_s^\infty(F) : P \to \mathbb{R}$ is given by

$$\alpha_s^\infty(F) = -\sum_{k=1}^\infty\sum_{j=1}^k \left(D_s\tilde{\boldsymbol{\mathcal{E}}}^{(k-j)\dagger}(-\partial_t)^{j-1}\frac{\delta\mathcal{K}^\infty}{\delta\partial_t^k\boldsymbol{E}}\right). \tag{4.41}$$

Here, $(-\partial_t)^{j-1}(\delta\mathcal{K}/\delta\partial_t^k\boldsymbol{E})$ must be expressed in terms of F by evaluating the functional derivative, applying the time derivative, and then using (4.36); and the adjoint operation, which is indicated by $\cdot^\dagger$, is defined by the following identity,

$$\int \delta\boldsymbol{E}\cdot D_s\tilde{\boldsymbol{\mathcal{E}}}^{(k-j)}[\delta F_s]\,\mathrm{d}^3\boldsymbol{X} = \int D_s\tilde{\boldsymbol{\mathcal{E}}}^{(k-j)\dagger}[\delta\boldsymbol{E}]\,\delta F_s\,\mathrm{d}\Omega_s. \tag{4.42}$$

In terms of the dressed Lagrange 1-form, the dressed Hamiltonian is

$$\tilde{\mathcal{H}}^\infty = \tilde{\vartheta}^\infty[\tilde{V}^\infty] - \tilde{L}^\infty, \tag{4.43}$$

where $\tilde{V}^\infty$ and $\tilde{L}^\infty$ are given by dressing equations (4.33) and (4.25) using (4.36). More explicitly, $\tilde{\mathcal{H}}^\infty$ is given by

$$\begin{aligned}\tilde{\mathcal{H}}(F) &= \sum_s \int (K_s^\infty(j\tilde{\boldsymbol{\mathcal{E}}}(F)) + \tilde{V}_s^{\infty i}\partial_{z^i}\alpha_s^\infty(F))\, F_s\,\mathrm{d}\Omega_s \\ &\quad + \frac{1}{2}\int \epsilon_o\tilde{\boldsymbol{\mathcal{E}}}\cdot\tilde{\boldsymbol{\mathcal{E}}}\,\mathrm{d}^3\boldsymbol{X} - \int\left(\sum_{k=0}^\infty(-\partial_t)^k\frac{\delta\mathcal{K}^\infty}{\delta\partial_t^k\boldsymbol{E}}\right)\cdot\tilde{\boldsymbol{\mathcal{E}}}\,\mathrm{d}^3\boldsymbol{X}.\end{aligned} \tag{4.44}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

The all-orders system phase-space Lagrangian is therefore

$$\mathcal{L}^{\infty}(F, V) = \tilde{\vartheta}^{\infty}[V] - \tilde{\mathcal{H}}^{\infty}(F). \tag{4.45}$$

When varying this Lagrangian, the now-familiar constraints,

$$\delta V_s = \dot{\xi}_s + L_{V_s}\xi_s \tag{4.46}$$

$$\delta F_s = -\frac{1}{\mathcal{J}_s}\mathrm{div}(\mathcal{J}_s\xi_s F_s), \tag{4.47}$$

must be imposed.

4.2. *The truncated system phase-space Lagrangian*

As it has been presented so far, the all-orders system phase-space Lagrangian is a purely formal result. However, I will now show that it can be exploited to build an attractive alternative theory to conventional first-order MSST drift kinetics. This alternative theory, which I will refer to as renormalized first-order MSST drift kinetics, will have the following desirable features: (1) a system phase space with the correct size, (2) first-order (in magnetic non-uniformity) accuracy, and (3) a Lagrangian formulation.

The procedure I will use to derive renormalized MSST drift kinetics mirrors the one I used earlier to derive renormalized two-oscillator drift kinetics. First the all-orders system phase-space Lagrangian must be expanded in powers of $\epsilon = \rho/L$, which is formally equivalent to expanding in powers of the inverse elementary charge, $|e|$ (note that $q_s = Z_s|e|$, where Z_s is some integer). For the sake of achieving first-order accuracy, both the dressed Lagrange 1-form and the dressed Hamiltonian will be expanded up to first order if $|e|^{-1}$.

In order to calculate $\tilde{\vartheta}$ and $\tilde{\mathcal{H}}$ with the desired level of accuracy, it is necessary to have explicit expression for some parts of the operator $j\tilde{\boldsymbol{\mathcal{E}}}(F)$. To be precise, if $\tilde{\boldsymbol{\mathcal{E}}}$ and $\tilde{\boldsymbol{\mathcal{E}}}^{(1)}$ are expanded in powers of ϵ as

$$\tilde{\boldsymbol{\mathcal{E}}} = \tilde{\boldsymbol{\mathcal{E}}}_o + \tilde{\boldsymbol{\mathcal{E}}}_1 + \cdots \tag{4.48}$$

$$\tilde{\boldsymbol{\mathcal{E}}}^{(1)} = \tilde{\boldsymbol{\mathcal{E}}}^{(1)}_o + \tilde{\boldsymbol{\mathcal{E}}}^{(1)}_1 + \cdots, \tag{4.49}$$

where subscripts denote order in ϵ, $\tilde{\boldsymbol{\mathcal{E}}}_o$, $\tilde{\boldsymbol{\mathcal{E}}}_1$, and $\tilde{\boldsymbol{\mathcal{E}}}^{(1)}_o$ must be known. One way to calculate these quantities is to find the first few orders in the asymptotic expansion of the operator $\mathcal{F}$ and then evaluate the composition $j\boldsymbol{\mathcal{E}} \circ \mathcal{F}$. However, the most direct approach is to analyse the Poisson equation in the all-orders MSST Euler–Lagrange equations, namely (4.34). It is straightforward to show that the dominant balance of (4.34) is given by

$$\boldsymbol{\nabla} \cdot (\epsilon_o \tilde{\boldsymbol{\epsilon}} \cdot \tilde{\boldsymbol{\mathcal{E}}}_o) = \rho(F), \tag{4.50}$$

where the dyad $\tilde{\boldsymbol{\epsilon}}$ is given by

$$\tilde{\boldsymbol{\epsilon}} = 1 + \left(\sum_s \frac{c^2\mu_o m_s N_s}{B^2}\right)(1 - \hat{b}\hat{b}). \tag{4.51}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

Because $\nabla \times \tilde{\boldsymbol{\mathcal{E}}}_o = 0$, this equation uniquely determines $\tilde{\boldsymbol{\mathcal{E}}}_o$. Indeed, if

$$G_\perp = (\Delta + \Delta_\perp)^{-1}, \tag{4.52}$$

where the negative-definite elliptic operator $\Delta + \Delta_\perp$ is given by $(\Delta + \Delta_\perp)\varphi = \nabla \cdot (\tilde{\epsilon} \cdot \nabla\varphi)$, (4.50) prescribes $\tilde{\boldsymbol{\mathcal{E}}}_o$ as

$$\tilde{\boldsymbol{\mathcal{E}}}_o(F) = \frac{1}{\epsilon_o} \nabla G_\perp[\rho(F)]. \tag{4.53}$$

Note that because the operator $G_\perp$ depends on F through the N_s in $\Delta_\perp$, $\tilde{\boldsymbol{\mathcal{E}}}_o$ is a nonlinear functional of F. Similarly, a dominant balance of the time derivative of (4.50) gives

$$\tilde{\boldsymbol{\mathcal{E}}}_o^{(1)}(F) = -\frac{1}{\epsilon_o} \nabla G_\perp[\nabla \cdot \mathrm{Im}(F)], \tag{4.54}$$

where

$$\begin{aligned} \mathrm{Im}(F) \;=\; & \sum_s \Big(q_s \langle \dot{\boldsymbol{X}}_{os} + \dot{\boldsymbol{X}}_{1s} \rangle_s N_s \\ & - \frac{m_s}{B^2} \nabla \cdot (\langle \dot{\boldsymbol{X}}_{os} \rangle_s N_s)(1 - \hat{b}\hat{b}) \cdot \tilde{\boldsymbol{\mathcal{E}}}_o \Big); \end{aligned} \tag{4.55}$$

the angle brackets denote velocity-space averages, $\langle \boldsymbol{Q} \rangle_s = N_s^{-1} \int \boldsymbol{Q} F_s \mathcal{J}_s \, \mathrm{d}v_\parallel \, \mathrm{d}\mu \, \mathrm{d}\zeta$; and

$$\dot{\boldsymbol{X}}_{os} = v_\parallel \hat{b} + \boldsymbol{u}_{Eo} \tag{4.56}$$

$$\begin{aligned} \dot{\boldsymbol{X}}_{1s} = & \left(\frac{u_{Eo}^2}{\omega_{cs}} \tau + \frac{2 v_\parallel}{\omega_{cs}} \hat{b} \times \boldsymbol{u}_{Eo} \cdot \kappa - \frac{\boldsymbol{u}_{Eo} \cdot \nabla \times \boldsymbol{u}_{Eo}}{2\,\omega_{cs}} \right) \hat{b} \\ & + \frac{\mu\, \hat{b} \times \nabla B}{q_s B} + \frac{v_\parallel^2 \hat{b} \times \kappa}{\omega_{cs}} - \frac{\hat{b} \times \nabla u_{Eo}^2}{2\,\omega_{cs}}, \end{aligned} \tag{4.57}$$

with $\boldsymbol{u}_{Eo} = \tilde{\boldsymbol{\mathcal{E}}}_o(F) \times \hat{b} B^{-1}$, are the $O(1)$ and $O(\epsilon)$ contributions to the guiding centre drift velocity, respectively. Next, $\tilde{\boldsymbol{\mathcal{E}}}_1$ is determined by the first sub-dominant contribution to the all-orders Poisson equation, namely

$$\nabla \cdot (\epsilon_o \tilde{\epsilon} \cdot \tilde{\boldsymbol{\mathcal{E}}}_1) = -\nabla \cdot \boldsymbol{\pi}_1(F), \tag{4.58}$$

where

$$\begin{aligned} \boldsymbol{\pi}_1(F) \;&=\; -\frac{\delta \mathcal{K}_1}{\delta \boldsymbol{E}} + \partial_t \frac{\delta \mathcal{K}_1}{\delta \dot{\boldsymbol{E}}} \\ &=\; -\sum_s \left(\frac{q_s N_s}{2\omega_{cs}^2} \dot{\boldsymbol{u}}_{Eo} - \frac{q_s}{2\omega_{cs}^2} \nabla \cdot (\langle \boldsymbol{X}_{os} \rangle_s N_s) \boldsymbol{u}_{Eo} \right) - \frac{\delta \mathcal{K}_1}{\delta \boldsymbol{E}}(\tilde{\boldsymbol{\mathcal{E}}}_o, \tilde{\boldsymbol{\mathcal{E}}}_o^{(1)}), \end{aligned} \tag{4.59}$$

$\mathcal{K}_1 = \sum_s \int K_{1s} F_s \, \mathrm{d}\Omega_s$, and $\dot{\boldsymbol{u}}_{Eo} = \tilde{\boldsymbol{\mathcal{E}}}_o^{(1)}(F) \times \hat{b} B^{-1}$. Therefore

$$\tilde{\boldsymbol{\mathcal{E}}}_1(F) = -\frac{1}{\epsilon_o} \nabla G_\perp[\nabla \cdot \boldsymbol{\pi}_1(F)]. \tag{4.60}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

Because doing so will lead to considerable simplifications of the renormalized drift kinetic theory later on, it is also useful to introduce the 'second-order electric field' $\tilde{\boldsymbol{\mathcal{E}}}_2(F)$, which is specified by the equation

$$\boldsymbol{\nabla}\cdot(\epsilon_o\tilde{\epsilon}\cdot\tilde{\boldsymbol{\mathcal{E}}}_2+\boldsymbol{\pi}(F)-\boldsymbol{\pi}_1(F))=0, \tag{4.61}$$

where

$$\boldsymbol{\pi}(F)=-\frac{\delta\mathcal{K}_1}{\delta\boldsymbol{E}}(\underline{\boldsymbol{\mathcal{E}}},\underline{\boldsymbol{\mathcal{E}}}^{(1)})+\frac{\delta^2\mathcal{K}_1}{\delta\boldsymbol{E}\delta\dot{\boldsymbol{E}}}[\underline{\boldsymbol{\mathcal{E}}}^{(1)}]+\frac{\delta^2\mathcal{K}_1}{\delta F\delta\dot{\boldsymbol{E}}}(\underline{\boldsymbol{\mathcal{E}}})[\dot{F}^*] \tag{4.62}$$

and

$$\dot{F}_s^*=-\mathcal{J}_s^{-1}\boldsymbol{\nabla}\cdot(\mathcal{J}_s(\dot{\boldsymbol{X}}_{os}+\dot{\boldsymbol{X}}_{1s})F_s) \tag{4.63}$$

$$\underline{\boldsymbol{\mathcal{E}}}=\tilde{\boldsymbol{\mathcal{E}}}_o+\tilde{\boldsymbol{\mathcal{E}}}_1 \tag{4.64}$$

$$\underline{\boldsymbol{\mathcal{E}}}^{(1)}=D\underline{\boldsymbol{\mathcal{E}}}[\dot{F}^*]. \tag{4.65}$$

It is straightforward to verify that the difference $\boldsymbol{\pi}-\boldsymbol{\pi}_1=O(\epsilon^2)$, which implies that $\tilde{\boldsymbol{\mathcal{E}}}_2=O(\epsilon^2)$.

With these expressions for $\tilde{\boldsymbol{\mathcal{E}}}_o$, $\tilde{\boldsymbol{\mathcal{E}}}_1$ and $\tilde{\boldsymbol{\mathcal{E}}}_o^{(1)}$ in hand, the truncated system phase-space Lagrangian can be found. The truncated dressed Lagrange 1-form is given by

$$\tilde{\vartheta}[\xi]=\sum_s\int\tilde{\ell}_s^{symp}(z,\xi_s(z);F)F_s\,\mathrm{d}\Omega_s, \tag{4.66}$$

where

$$\tilde{\ell}_s^{symp}(z,\dot{z};F)=\ell_s^{symp}(z,\dot{z})+\dot{\boldsymbol{X}}\cdot\boldsymbol{\nabla}\alpha_s, \tag{4.67}$$

the function α_s is given by

$$\begin{aligned}\alpha_s &= -D_s\tilde{\boldsymbol{\mathcal{E}}}_o^{\dagger}\frac{\delta\mathcal{K}}{\delta\dot{\boldsymbol{E}}}\\ &= D_s\tilde{\boldsymbol{\mathcal{E}}}_o^{\dagger}\left[\sum_{\bar{s}}\frac{m_sN_s}{2B^2\omega_{cs}}\tilde{\boldsymbol{\mathcal{E}}}_o\times\hat{b}\right],\end{aligned} \tag{4.68}$$

and $\mathcal{K}=\mathcal{K}_o+\mathcal{K}_1=\sum_s\int(K_{os}+K_{1s})F_s\,\mathrm{d}\Omega_s$. The truncated dressed Hamiltonian is given by

$$\tilde{\mathcal{H}}(F)=\sum_s\int(\tilde{K}_s(F)F_s+\alpha_s^*\dot{F}_s^*)\,\mathrm{d}\Omega_s+\frac{1}{2}\int\epsilon_o\tilde{\boldsymbol{\mathcal{E}}}\cdot\tilde{\boldsymbol{\mathcal{E}}}\,\mathrm{d}^3\boldsymbol{X}+\int\tilde{\boldsymbol{P}}\cdot\tilde{\boldsymbol{\mathcal{E}}}\,\mathrm{d}^3\boldsymbol{X}, \tag{4.69}$$

where $\tilde{\boldsymbol{\mathcal{E}}}=\tilde{\boldsymbol{\mathcal{E}}}_o+\tilde{\boldsymbol{\mathcal{E}}}_1+\tilde{\boldsymbol{\mathcal{E}}}_2$;

$$\tilde{K}_s(F)=K_{os}(\tilde{\boldsymbol{\mathcal{E}}})+K_{1s}(\underline{\boldsymbol{\mathcal{E}}},\underline{\boldsymbol{\mathcal{E}}}^{(1)}); \tag{4.70}$$

$$\alpha_s^*=-D_s\underline{\boldsymbol{\mathcal{E}}}^{\dagger}\frac{\delta\mathcal{K}_1}{\delta\dot{\boldsymbol{E}}}(\underline{\boldsymbol{\mathcal{E}}}); \tag{4.71}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

and the polarization density $\tilde{\boldsymbol{P}}$ is given by

$$\tilde{\boldsymbol{P}} = -\frac{\delta \mathcal{K}_o}{\delta \boldsymbol{E}}(\tilde{\boldsymbol{\mathcal{E}}}) + \boldsymbol{\pi}(F). \tag{4.72}$$

Note that $\boldsymbol{\nabla} \cdot (\epsilon_o \tilde{\boldsymbol{\mathcal{E}}} + \tilde{\boldsymbol{P}}) = \rho(F)$, which can be carefully applied to show that the functional derivative of $\tilde{\mathcal{H}}$ is given by

$$\frac{\delta \tilde{\mathcal{H}}}{\delta F_s} = H_s + \sum_{\bar{s}} \eta^*_{s\bar{s}}[\dot{F}^*_{\bar{s}}] + D_s \tilde{\boldsymbol{\mathcal{E}}}^{\dagger}_2[\boldsymbol{\pi}], \tag{4.73}$$

where $H_s = q_s \tilde{\varphi} + \tilde{K}_s$, $-\boldsymbol{\nabla}\tilde{\varphi} = \tilde{\boldsymbol{\mathcal{E}}}$, and $\eta^*_{s\bar{s}} = D_s \alpha^{*\dagger}_{\bar{s}} - D_{\bar{s}} \alpha^*_s$. The truncated system phase-space Lagrangian is therefore

$$\mathcal{L}(F, V) = \tilde{\vartheta}[V] - \tilde{\mathcal{H}}. \tag{4.74}$$

Upon applying the constrained variations

$$\delta V_s = \dot{\xi}_s + L_{V_s} \xi_s \tag{4.75}$$

$$\delta F_s = -\frac{1}{\mathcal{J}_s} \mathrm{div}(\mathcal{J}_s \xi_s F_s) \tag{4.76}$$

to this Lagrangian, the Euler–Lagrange equations that result are given by

$$V^i_s = \left\{ z^i, \varepsilon_s + \frac{\delta \tilde{\mathcal{H}}}{\delta F_s} \right\}^{gc}_s, \tag{4.77}$$

where

$$\begin{aligned} \varepsilon_s &= \sum_{\bar{s}} (D_s \alpha^{\dagger}_{\bar{s}} - D_{\bar{s}} \alpha_s)[\mathcal{J}^{-1}_{\bar{s}} \mathrm{div}(\mathcal{J}_{\bar{s}} V_{\bar{s}} F_{\bar{s}})] \\ &\equiv \sum_{\bar{s}} \eta_{s\bar{s}}[\mathcal{J}^{-1}_{\bar{s}} \mathrm{div}(\mathcal{J}_{\bar{s}} V_{\bar{s}} F_{\bar{s}})]. \end{aligned} \tag{4.78}$$

Note that because ε_s depends on V, (4.77) is an implicit equation for V. In order to solve for V, first note that

$$\begin{aligned} \varepsilon_s &= \sum_{\bar{s}} \eta_{s\bar{s}}[\{F_{\bar{s}}, \varepsilon_{\bar{s}}\}^{gc}_{\bar{s}} + \{F_{\bar{s}}, \delta \tilde{\mathcal{H}}/\delta F_{\bar{s}}\}^{gc}_{\bar{s}}] \\ &\equiv \sum_{\bar{s}} \chi_{s\bar{s}}[\varepsilon_{\bar{s}} + \delta \tilde{\mathcal{H}}/\delta F_{\bar{s}}]. \end{aligned} \tag{4.79}$$

Therefore,

$$\mathcal{D}[\varepsilon] = \chi[\delta \tilde{\mathcal{H}}/\delta F], \tag{4.80}$$

where the matrix of operators $\mathcal{D}$ is given by

$$(\mathcal{D}[\varepsilon])_s = \sum_{\bar{s}} \mathcal{D}_{s\bar{s}}[\varepsilon_{\bar{s}}] = \sum_{\bar{s}} (\delta_{s\bar{s}} - \chi_{s\bar{s}})[\varepsilon_{\bar{s}}]. \tag{4.81}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

In appendix B, I argue that $\mathcal{D}$ should be invertible as a mapping from multi-species phase-space functions into itself when ϵ is small. Thus,

$$\varepsilon = \mathcal{D}^{-1}\chi[\delta\tilde{\mathcal{H}}/\delta F], \tag{4.82}$$

which, in conjunction with (4.77), determines the Eulerian phase-space fluid velocity V explicitly,

$$\begin{aligned} V^i &= \left\{ z^i, \mathcal{D}^{-1}\chi\left[\frac{\delta\tilde{\mathcal{H}}}{\delta F}\right] + \frac{\delta\tilde{\mathcal{H}}}{\delta F} \right\}^{gc} \\ &= \left\{ z^i, \mathcal{D}^{-1}\frac{\delta\tilde{\mathcal{H}}}{\delta F} \right\}^{gc}. \end{aligned} \tag{4.83}$$

The time evolution equation for F in renormalized MSST drift kinetics is therefore

$$\partial_t F = -\left\{ F, \mathcal{D}^{-1}\frac{\delta\tilde{\mathcal{H}}}{\delta F} \right\}^{gc}, \tag{4.84}$$

which is manifestly a first-order ODE on F-space. Recall that the functional derivative $\delta\tilde{\mathcal{H}}/\delta F$ is given in (4.73).

I will conclude this section by analysing the linearization of renormalized MSST drift kinetics about the same type of slab equilibria used in § 3. Again I will assume the background magnetic field is uniform and z-directed. It is straightforward to show that when F_s is given by (3.15) with $N_{os} = n(1 + x/L)$, $\tilde{\boldsymbol{\mathcal{E}}}_o = \tilde{\boldsymbol{\mathcal{E}}}_1 = \tilde{\boldsymbol{\mathcal{E}}}_o^{(1)} = 0$, the functional derivatives $\delta\tilde{\mathcal{H}}/\delta F_s = (1/2)m_s v_\parallel^2 + \mu B$, and $\chi_{s\bar{s}}[\delta\tilde{\mathcal{H}}/\delta F_{\bar{s}}] = 0$. Therefore the equilibria from § 3 are indeed equilibria of renormalized MSST drift kinetics. When renormalized MSST drift kinetics is linearized about this type of equilibrium, the dynamical equations governing z-independent fluctuations reduce to the following pair of linear integro-differential equations,

$$-\epsilon_o(\Delta + \Delta_\perp)\delta\dot{\varphi}_1 = \frac{\alpha}{2}\nabla_\perp^2\partial_y\delta\varphi_A + \gamma\nabla_\perp^4\partial_y\delta\varphi_B \tag{4.85}$$

$$-\epsilon_o(\Delta + \Delta_\perp)\delta\dot{\varphi}_o = \frac{\alpha}{2}\nabla_\perp^2\partial_y\delta\varphi_B, \tag{4.86}$$

where φ_o and φ_1 are the $0th$ and $1st$ order contributions to the electrostatic potential, and

$$\begin{aligned} \delta\varphi_A = \delta\varphi_o + \delta\varphi_1 + \frac{\alpha\beta}{2}\frac{G_\perp}{\epsilon_o}\partial_y^2\nabla_\perp^2\frac{G_\perp}{\epsilon_o}(\delta\varphi_1 + \delta\varphi_2) \\ - \frac{\alpha\beta}{2}\left(\frac{G_\perp}{\epsilon_o}\right)^2\nabla_\perp^2\partial_y^2(\delta\varphi_o + \delta\varphi_1 - \delta\varphi_2) \end{aligned} \tag{4.87}$$

$$\delta\varphi_B = \delta\varphi_o + \delta\varphi_1 - \delta\varphi_2 \tag{4.88}$$

$$\delta\varphi_2 = -\frac{\alpha\beta}{2}\left(\frac{G_\perp}{\epsilon_o}\right)^2\nabla_\perp^2\partial_y^2\delta\varphi_o. \tag{4.89}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

Applying the same WKB analysis to this system of equations as in § 3 shows that there are two normal modes with $k_x = 0$. The two branches of the dispersion relation are given by

$$\omega_d = -\frac{v_* k_y}{1 + \frac{x}{L} + \frac{v_{Ao}^2}{c^2}} + O(\epsilon^2) \tag{4.90}$$

$$\omega_s = O(\epsilon^2). \tag{4.91}$$

The first branch, ω_d, recovers the usual drift wave (with a slightly modified frequency), while the second branch represents an artificial very low-frequency wave. The renormalization procedure has therefore managed to eliminate the artificial high-frequency wave, $\omega_\times$, from MSST drift kinetics and replace it with the artificial low-frequency wave, ω_s. From a computational point of view, an artificial low-frequency wave may be preferable to an artificial high-frequency wave. For instance, in an explicit integration scheme, the time-step restriction imposed by a high frequency wave would be more severe than the time-step restriction imposed by a low-frequency wave.

5. Discussion

One approach to dealing with higher-order time derivatives of the electrostatic potential that appear in the equations defining higher-order drift kinetic and gyrokinetic theories would be to first neglect the terms involving the time derivatives, solve for the potential, then use this approximate potential to evaluate the neglected time derivative. Such a manipulation, which can be conceptualized as identifying the so-called slow manifold (Jones 1995), is not obviously consistent with energy and momentum conservation. One way of interpreting the results of this paper is that they show how this 'bootstrapping' technique for dealing with higher time derivatives of the potential can be modified to be fully consistent with Lagrangian mechanics, and therefore energy and momentum conservation. Another approach to dealing with extra time derivatives of the electromagnetic field has been developed in the context of a drift kinetic formalism that makes use of neither near-identity phase-space transformations nor Lagrangian mechanics (Ramos 2008). There the time derivatives of the electromagnetic field are eliminated through the use of exact momentum conservation relations. In the end, the technique in Ramos (2008) produces a hybrid fluid-drift kinetic system with exact conservation laws for energy and momentum, and without an apparent variational principle. It would be interesting to learn whether or not this hybrid system could be cast in the framework of Hamiltonian (Tronci 2010) or Lagrangian (Holm & Tronci 2012) hybrid models.

While I have been stressing the fact that renormalized MSST drift kinetics satisfies a variational principle, renormalized MSST drift kinetics is also an infinite-dimensional Hamiltonian system. The Hamiltonian functional is given by $\tilde{\mathcal{H}}$ and the field-theoretic Poisson bracket is given by

$$[\mathcal{F}, \mathcal{G}] = \int \left\{ \mathcal{D}^{-1} \frac{\delta \mathcal{F}}{\delta F}, \mathcal{D}^{-1} \frac{\delta \mathcal{G}}{\delta F} \right\}^{gc} F \, \mathrm{d}\Omega + \int \eta \left[\left\{ F, \mathcal{D}^{-1} \frac{\delta \mathcal{F}}{\delta F} \right\}^{gc} \right] \left\{ F, \mathcal{D}^{-1} \frac{\delta \mathcal{G}}{\delta F} \right\}^{gc} \mathrm{d}\Omega, \tag{5.1}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

where $\mathcal{F}, \mathcal{G}$ are arbitrary functions of the multi-species distribution function, and the skew-symmetric operator η is defined in (4.78). This Poisson bracket can be deduced directly from the phase-space Lagrangian for renormalized MSST drift kinetics following the example set in Burby (2015), and therefore automatically satisfies the Jacobi identity. This means that, aside from eliminating unphysical rapidly varying modes, an additional benefit of employing renormalized MSST drift kinetics instead of conventional MSST drift kinetics is that the dynamically accessible stability method (Morrison & Pfirsch 1989; Burby *et al.* 2015) can be applied to the renormalized theory in order to assess the stability of equilibria. In contrast, conventional MSST drift kinetics has no known Hamiltonian structure. Note that this is consistent with the existence of a Hamiltonian formulation for the low-order electrostatic gyrokinetic system given in Squire *et al.* (2013).

The appearance of higher time derivatives of φ in MSST drift kinetics can be traced back to the appearance of higher time derivatives in the conventional drift kinetic system Lagrangian. In the theory of Lagrangian systems with non-degenerate Lagrangians, higher time derivatives in the Lagrangian necessarily indicate the presence of the so-called Ostrogradsky instability (Ostrogradsky 1850; Motohashi & Suyama 2015). In light of this well-known phenomenon, it is perhaps not entirely surprising that there are unphysical modes present in MSST drift kinetics. However, because the MSST system Lagrangian is degenerate, it is not completely justified to invoke Ostrogradsky's theorem as an *a priori* explanation of the unphysical modes in MSST drift kinetics. It would nevertheless be interesting to attempt to extend Ostrogradsky's theorem to allow for Lagrangians with the particular type of degeneracy inherent in the MSST system Lagrangian.

Looking ahead, two important extensions of the work presented here are (1) incorporating electromagnetic effects into the renormalization procedure, and (2) identifying a collision operator with exact conservation properties for renormalized MSST drift kinetics. The second task may prove to be especially challenging because it seems likely that, due to the use of the 'bootstrapping' technique, collisional renormalized MSST drift kinetics will contain the collision operator in both the kinetic equation and the drift kinetic Poisson equation. Nevertheless, it would be interesting to know if results akin to those given in Burby & Brizard (2015) exist for renormalized MSST drift kinetics.

Finally, it is worth remarking on the relevance of the methods of this paper to gyrokinetic modelling in the tokamak edge. Due to the particular orderings relevant in the edge (Scott 2016), the time derivatives that appear in the gyrocentre Hamiltonian are likely to be too small to be dynamically relevant. Therefore, the results in this paper are unlikely to be important to gyrokinetic modelling of the edge.

Acknowledgements

I would like to acknowledge helpful discussions with I. Abel and the magneto-fluid dynamics group at the Courant Institute. This work was supported by the US Department of Energy, Office of Science, Fusion Energy Sciences under award no. DE-FG02-86ER53223.

Appendix A. Why does 'dressing' work?

At two points in the main text I applied a dressing procedure to derive a phase-space Lagrangian for a Lagrangian system. The purpose of this appendix is to give a precise definition of the term 'phase space Lagrangian,' and provide a

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

rigorous justification for using the 'dressing' procedure to find a Lagrangian system's phase-space Lagrangian. For the sake of making the discussion accessible, I will strive to convey these general ideas by way of a specific example, namely all-orders two-oscillator drift kinetics. For a similar discussion in the context of variational kinetic theories, see Burby (2015). For a discussion of phase-space Lagrangians in the context of magnetic field line flow, see Cary & Littlejohn (1983).

Suppose $\mathcal{P}$ is a linear space with coordinates z^i and that X is a vector field on $\mathcal{P}$. The equation $\dot{z} = X(z)$, where $z \in \mathcal{P}$, specifies an autonomous system on $\mathcal{P}$ that will be referred to as X. The phase space for the system X is the space $\mathcal{P}$. X admits a phase-space Lagrangian if there is a 1-form on $\mathcal{P}$, $\vartheta = \vartheta_i \boldsymbol{d} z^i$, and a function on $\mathcal{P}$, h, such that

$$\iota_X \boldsymbol{d}\vartheta = -\boldsymbol{d}h \Leftrightarrow (\vartheta_{i,j} - \vartheta_{j,i})X^j = -h_{,i}. \tag{A 1}$$

The 1-form ϑ and the function h are X's Lagrange 1-form and Hamiltonian function, respectively. Given an X that admits a phase-space Lagrangian, X's phase-space Lagrangian, ℓ, is the function on $\mathcal{P} \times \mathcal{P}$ given by

$$\ell(z, \dot{z}) = \vartheta_i(z)\dot{z}^i - h(z). \tag{A 2}$$

The justification for this terminology is the following. Given an X that admits a phase-space Lagrangian, ℓ, define the functional $S : \mathfrak{P}(P) \to \mathbb{R}$, where $\mathfrak{P}(P)$ is the space of paths, γ, through $\mathcal{P}$ parameterized by the fixed interval $[t_1, t_2]$, using the formula

$$\begin{aligned} S(\gamma) &= \int_\gamma \vartheta - \int_{t_1}^{t_2} h(\gamma(t))\,\mathrm{d}t \\ &= \int_{t_1}^{t_2} \ell(\gamma(t), \dot{\gamma}(t))\,\mathrm{d}t. \end{aligned} \tag{A 3}$$

If γ is an integral curve of X, i.e. $\dot{\gamma}(t) = X(\gamma(t))$ for each $t \in [t_1, t_2]$, then the first fixed-endpoint variation of S is zero at γ. Thus, the autonomous system X satisfies a variational principle, and ℓ serves as the corresponding Lagrangian function. The modifier 'phase space' in 'phase-space Lagrangian' corresponds to the fact that the functional S takes as its argument a path in phase space; the most familiar variational principle, namely Hamilton's principle, is phrased in terms of a functional whose argument is a path in configuration space.

A natural class of systems that admit a phase-space Lagrangian consists of all systems that satisfy variational principles. I will illustrate this point in the context of all-orders two-oscillator drift kinetics. In this case, $\mathcal{P}$ is $(Q, P, \varphi, \dot{\varphi})$-space, and the vector field X is given by

$$X^Q = (1 - 2\epsilon g\hat{q}_-(Q, \varphi))P - \epsilon\omega_2^2\dot{\varphi} \tag{A 4}$$

$$X^P = -(\omega_1^2 + 2g\varphi)\hat{q}_-(Q, \varphi) \tag{A 5}$$

$$X^\varphi = \dot{\varphi} \tag{A 6}$$

$$X^{\dot{\varphi}} = -\omega_2^2\varphi - g\hat{q}_-^2(Q, \varphi). \tag{A 7}$$

This system admits the following variational principle. Let the set $\mathcal{Q}$ be (Q, P, φ)-space and let $\mathfrak{P}(\mathcal{Q})$ be the corresponding space of paths, $c = (Q, P, \varphi)$, through $\mathcal{Q}$ parameterized by $[t_1, t_2]$. Define the functional $S_o : \mathfrak{P}(\mathcal{Q}) \to \mathbb{R}$ using the formula

$$S_o(c) = \int_{t_1}^{t_2} \bar{L}(c(t), \dot{c}(t), \ddot{\varphi}(t), \dddot{\varphi}(t))\,\mathrm{d}t, \tag{A 8}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

where $\bar{L}$ is given in (2.24). If the first fixed-endpoint variation of S_o vanishes at $c = (Q, P, \varphi)$, then, as described in § 2, $\gamma = (Q, P, \varphi, \dot{\varphi})$ is an integral curve of X. Conversely, if $\gamma = (Q, P, \varphi, \dot{\varphi})$ is an integral curve of X, then the first fixed-endpoint variation of S_o vanishes at $c = (Q, P, \varphi)$. Note that $\bar{L}$ is not a phase-space Lagrangian, i.e. it is not of the form expressed in (A 2). Nevertheless, X does admit a phase-space Lagrangian. To see this, first let $F_t : \mathcal{P} \to \mathcal{P}$ be the flow map of the vector field X; F_t is the unique one-parameter family of maps $\mathcal{P} \to \mathcal{P}$ that satisfies $F_0(z) = z$ and $\partial_t F_t(z) = X(F_t(z))$ for each $z \in \mathcal{P}$. Using F_t, it is possible to construct a mapping $\mathrm{sol} : \mathcal{P} \to \mathfrak{P}(\mathcal{P})$ according to the formula

$$\mathrm{sol}(z)(t) = F_{t-t_1}(z). \tag{A 9}$$

Note that $\mathrm{sol}(z)$ is the unique path through $\mathcal{P}$ that (1) is an integral curve of the vector field X and (2) passes through the point z at $t = t_1$. Because any path $\gamma = (Q, P, \varphi, \dot{\varphi})$ through $\mathcal{P}$ determines a corresponding path, $\pi(\gamma) = (Q, P, \varphi)$, through $\mathcal{Q}$, sol and S_o can be used to construct the function on phase space $\bar{S} = S_o \circ \pi \circ \mathrm{sol} : \mathcal{P} \to \mathbb{R}$. $\bar{S}$ may be thought of as the action S_o regarded as a function of initial conditions. The phase-space Lagrangian for all-orders two-oscillator drift kinetics can be derived by analysing the derivatives of $\bar{S}$ with respect to z and t_2. ($\bar{S}$ depends on t_2 because the upper limit of the time integral in the definition of S_o is t_2.)

In order to calculate the derivatives of $\bar{S}$, it will be convenient to introduce the curves $\gamma_\epsilon = \mathrm{sol}(z + \epsilon v) = (Q_\epsilon, P_\epsilon, \varphi_\epsilon, \dot{\varphi}_\epsilon)$, $c_\epsilon = \pi(\gamma_\epsilon)$, Γ_ϵ, and Υ_ϵ, where ϵ is a small positive real number, $v \in \mathcal{P}$ is an arbitrary four-component vector, and

$$\Gamma_\epsilon(t) = \left(Q_\epsilon(t), P_\epsilon(t), \varphi_\epsilon(t), \frac{\mathrm{d}Q_\epsilon}{\mathrm{d}t}, \frac{\mathrm{d}P_\epsilon}{\mathrm{d}t}, \frac{\mathrm{d}\varphi_\epsilon}{\mathrm{d}t}, \frac{\mathrm{d}^2\varphi_\epsilon}{\mathrm{d}t^2}, \frac{\mathrm{d}^3\varphi_\epsilon}{\mathrm{d}t^3}\right) \tag{A 10}$$

$$\Upsilon_\epsilon(t) = \left(Q_\epsilon(t), P_\epsilon(t), \varphi_\epsilon(t), \frac{\mathrm{d}Q_\epsilon}{\mathrm{d}t}, \frac{\mathrm{d}P_\epsilon}{\mathrm{d}t}, \frac{\mathrm{d}^2P_\epsilon}{\mathrm{d}t^2}, \frac{\mathrm{d}\varphi_\epsilon}{\mathrm{d}t}, \frac{\mathrm{d}^2\varphi_\epsilon}{\mathrm{d}t^2}, \frac{\mathrm{d}^3\varphi_\epsilon}{\mathrm{d}t^3}\right). \tag{A 11}$$

Note that because γ_ϵ is an integral curve of X for each ϵ, it must be true that

$$\frac{\mathrm{d}Q_\epsilon}{\mathrm{d}t} = X^Q(\gamma_\epsilon(t)) \tag{A 12}$$

$$\frac{\mathrm{d}P_\epsilon}{\mathrm{d}t} = X^P(\gamma_\epsilon(t)) \tag{A 13}$$

$$\frac{\mathrm{d}^2P}{\mathrm{d}t^2} = (X^Q\partial_Q X^P + X^\varphi \partial_\varphi X^P)(\gamma_\epsilon(t)) \tag{A 14}$$

$$\frac{\mathrm{d}\varphi_\epsilon}{\mathrm{d}t} = X^\varphi(\gamma_\epsilon(t)) \tag{A 15}$$

$$\frac{\mathrm{d}^2\varphi_\epsilon}{\mathrm{d}t^2} = \frac{\mathrm{d}\dot{\varphi}_\epsilon}{\mathrm{d}t} = X^{\dot{\varphi}}(\gamma_\epsilon(t)) \tag{A 16}$$

$$\frac{\mathrm{d}^3\varphi_\epsilon}{\mathrm{d}t^3} = (X^Q\partial_Q X^{\dot{\varphi}} + X^\varphi \partial_\varphi X^{\dot{\varphi}})(\gamma_\epsilon(t)). \tag{A 17}$$

The directional derivative of $\bar{S}$ at z along v is given by

$$\begin{aligned} \boldsymbol{d}\bar{S}_z(v) &= \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_0 \bar{S}(z + \epsilon v) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_0 S_o(c_\epsilon), \end{aligned} \tag{A 18}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

which is equal to a certain free-endpoint variation of S_o at a curve ($c = c_0$) that satisfies the Euler–Lagrange equations. Applying integration by parts several times and making use of the fact that c_0 satisfies the Euler–Lagrange equations gives

$$\boldsymbol{d}\bar{S}_z(v) = [\vartheta_Q(\Upsilon_0(t))\delta Q(t) + \vartheta_\varphi(\Upsilon_0(t))\delta\varphi(t) + \vartheta_{\dot\varphi}(\Upsilon_0(t))\delta\dot\varphi(t) + \vartheta_{\ddot\varphi}(\Upsilon_0(t))\delta\ddot\varphi(t)]_{t_1}^{t_2}, \tag{A 19}$$

where

$$\delta Q(t) = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_0 Q_\epsilon(t) \tag{A 20}$$

$$\delta\varphi(t) = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_0 \varphi_\epsilon(t) \tag{A 21}$$

$$\delta\dot\varphi(t) = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_0 \frac{\mathrm{d}\varphi_\epsilon}{\mathrm{d}t} \tag{A 22}$$

$$\delta\ddot\varphi(t) = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_0 \frac{\mathrm{d}^2\varphi_\epsilon}{\mathrm{d}t^2}, \tag{A 23}$$

$$\vartheta_Q(Z) = \frac{\partial\bar{L}}{\partial\dot{Q}} \tag{A 24}$$

$$\vartheta_\varphi(Z) = \frac{\partial\bar{L}}{\partial\dot\varphi} - \frac{\partial^2\bar{L}}{\partial Q\partial\ddot\varphi}\dot{Q} - \frac{\partial^2\bar{L}}{\partial\varphi\partial\ddot\varphi}\dot\varphi - \frac{\partial^2\bar{L}}{\partial\ddot\varphi^2}\dddot\varphi + \frac{\partial^2\bar{L}}{\partial P\partial\dddot\varphi}\ddot{P} \tag{A 25}$$

$$\vartheta_{\dot\varphi}(Z) = \frac{\partial\bar{L}}{\partial\ddot\varphi} - \frac{\partial^2\bar{L}}{\partial P\partial\dddot\varphi}\dot{P} \tag{A 26}$$

$$\vartheta_{\ddot\varphi}(Z) = \frac{\partial\bar{L}}{\partial\dddot\varphi}, \tag{A 27}$$

and all partial derivatives are evaluated at $Z = (Q, P, \varphi, \dot{Q}, \dot{P}, \ddot{P}, \dot\varphi, \ddot\varphi, \dddot\varphi)$. The 1-form on Z-space $\vartheta = \vartheta_Q\,\boldsymbol{d}Q + \vartheta_\varphi\,\boldsymbol{d}\varphi + \vartheta_{\dot\varphi}\,\boldsymbol{d}\dot\varphi + \vartheta_{\ddot\varphi}\,\boldsymbol{d}\ddot\varphi$ is the bare Lagrange 1-form introduced in § 2. In terms of the bare Lagrange 1-form, the directional derivative of $\bar{S}$ is

$$\boldsymbol{d}\bar{S}_z(v) = [\vartheta_{\Upsilon_0(t)}(\delta\Upsilon_0(t))]_{t_1}^{t_2}. \tag{A 28}$$

This expression can be simplified further by noting that, according to (A 12)–(A 17), $\Upsilon_0(t) = \boldsymbol{J}(\gamma_0(t))$ and $\delta\Upsilon_0(t) = D\boldsymbol{J}_{\gamma_0(t)}[\delta\gamma(t)]$, where $\boldsymbol{J}$ maps phase space, $\mathcal{P}$, into Z-space according to $\boldsymbol{J} = (Q_{\boldsymbol{J}}, P_{\boldsymbol{J}}, \varphi_{\boldsymbol{J}}, \dot{Q}_{\boldsymbol{J}}, \dot{P}_{\boldsymbol{J}}, \ddot{P}_{\boldsymbol{J}}, \dot\varphi_{\boldsymbol{J}}, \ddot\varphi_{\boldsymbol{J}}, \dddot\varphi_{\boldsymbol{J}})$, with

$$Q_{\boldsymbol{J}}(z) = Q \tag{A 29}$$

$$P_{\boldsymbol{J}}(z) = P \tag{A 30}$$

$$\varphi_{\boldsymbol{J}}(z) = \varphi \tag{A 31}$$

$$\dot{Q}_{\boldsymbol{J}}(z) = X^Q(z) \tag{A 32}$$

$$\dot{P}_{\boldsymbol{J}}(z) = X^P(z) \tag{A 33}$$

$$\ddot{P}_{\boldsymbol{J}}(z) = (X^Q\partial_Q X^P + X^\varphi\partial_\varphi X^P)(z) \tag{A 34}$$

$$\dot\varphi_{\boldsymbol{J}}(z) = X^\varphi(z) \tag{A 35}$$

$$\ddot\varphi_{\boldsymbol{J}}(z) = X^{\dot\varphi}(z) \tag{A 36}$$

$$\dddot\varphi_{\boldsymbol{J}}(z) = (X^Q\partial_Q X^{\dot\varphi} + X^\varphi\partial_\varphi X^{\dot\varphi})(z). \tag{A 37}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

Thus,

$$\begin{aligned}\vartheta_{\Upsilon_0(t)}(\delta\Upsilon_0(t)) &= \vartheta_{\boldsymbol{J}(\gamma_0(t))}(D\boldsymbol{J}_{\gamma_0(t)}[\delta\gamma(t)]) \\ &= (\boldsymbol{J}^*\vartheta)_{\gamma_0(t)}(\delta\gamma(t)) \\ &\equiv \tilde{\vartheta}^{\infty}_{\gamma_0(t)}(\delta\gamma(t)),\end{aligned} \tag{A 38}$$

where $\tilde{\vartheta}^{\infty} = \boldsymbol{J}^*\vartheta$ denotes the pullback of the 1-form, ϑ, on Z-space to phase space along the mapping $\boldsymbol{J}$, and

$$\boldsymbol{d}\bar{S}_z(v) = [\tilde{\vartheta}_{\gamma_0(t)}(\delta\gamma(t))]^{t_2}_{t_1}. \tag{A 39}$$

A straightforward application of the definition of pullback (Abraham & Marsden 2008) shows that $\boldsymbol{J}^*\vartheta$ is precisely the dressed Lagrange 1-form introduced in § 2. Finally, using the fact that

$$\begin{aligned}\delta\gamma(t) &= \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_0 F_{t-t_1}(z+\epsilon v) \\ &= (DF_{t-t_1})_z[v],\end{aligned} \tag{A 40}$$

and applying the definition of pullback once more, the directional derivative of $\bar{S}$ can be expressed as

$$\boldsymbol{d}\bar{S}_z(v) = (F^*_{t_2-t_1}\tilde{\vartheta}^{\infty} - \tilde{\vartheta}^{\infty})_z(v), \tag{A 41}$$

i.e. $\boldsymbol{d}\bar{S} = F^*_{t_2-t_1}\tilde{\vartheta}^{\infty} - \tilde{\vartheta}^{\infty}$.

In order to find the t_2 derivative of $\bar{S}$, first note that

$$\begin{aligned}\bar{S}(z) &= \int_{t_1}^{t_2} \bar{L}(\Gamma_0(t))\,\mathrm{d}t \\ &= \int_{t_1}^{t_2} (\boldsymbol{j}^*\bar{L})(\gamma_0(t))\,\mathrm{d}t \\ &= \int_{t_1}^{t_2} (\boldsymbol{j}^*\bar{L})(F_{t-t_1}(z))\,\mathrm{d}t,\end{aligned} \tag{A 42}$$

where the mapping $\boldsymbol{j} = (Q_{\boldsymbol{J}}, P_{\boldsymbol{J}}, \varphi_{\boldsymbol{J}}, \dot{Q}_{\boldsymbol{J}}, \dot{P}_{\boldsymbol{J}}, \dot{\varphi}_{\boldsymbol{J}}, \ddot{\varphi}_{\boldsymbol{J}}, \dddot{\varphi}_{\boldsymbol{J}})$. Therefore,

$$\frac{\mathrm{d}}{\mathrm{d}t_2}\bar{S}(z) = F^*_{t_2-t_1}(\boldsymbol{j}^*\bar{L})(z). \tag{A 43}$$

When (A 41) is combined with (A 43), a straightforward application of exterior calculus identities given in Abraham & Marsden (2008) shows that X admits a phase-space Lagrangian. Indeed,

$$\begin{aligned}&\boldsymbol{d}\frac{\mathrm{d}\bar{S}}{\mathrm{d}t_2} = \frac{\mathrm{d}}{\mathrm{d}t_2}(F^*_{t_2-t_1}\tilde{\vartheta}^{\infty} - \tilde{\vartheta}^{\infty}) \\ &\Rightarrow F^*_{t_2-t_1}(\boldsymbol{d}\boldsymbol{j}^*\bar{L}) = F^*_{t_2-t_1}(L_X\tilde{\vartheta}^{\infty}) \\ &\Rightarrow -\boldsymbol{d}(\tilde{\vartheta}^{\infty}(X) - \boldsymbol{j}^*\bar{L}) = \iota_X\boldsymbol{d}\tilde{\vartheta}^{\infty},\end{aligned} \tag{A 44}$$

which is precisely the condition given in (A 1). X's Lagrange 1-form is given by $\tilde{\vartheta}^{\infty}$, while its Hamiltonian is given by $h = \tilde{\vartheta}^{\infty}(X) - \boldsymbol{j}^*\bar{L}$. Moreover, a careful analysis of the definitions of $\tilde{\vartheta}^{\infty}$ and h shows the dressing procedure described in § 2 correctly identifies the phase-space Lagrangian for all-orders two-oscillator drift kinetics.

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

Appendix B. Why is the operator $\mathcal{D}$ invertible?

The linear operator $\mathcal{D}$, which maps multi-species phase-space functions into multi-species phase-space functions, is defined by the expression

$$\mathcal{D}_{s\bar{s}}[\varepsilon_{\bar{s}}] = (\delta_{s\bar{s}} - \chi_{s\bar{s}})[\varepsilon_{\bar{s}}], \tag{B 1}$$

where the operator $\chi_{s\bar{s}}$ is given by

$$\begin{aligned}\chi_{s\bar{s}}[\varepsilon_{\bar{s}}] &= \eta_{s\bar{s}}[\{F_{\bar{s}}, \varepsilon_{\bar{s}}\}^{gc}_{\bar{s}}] \\ &= 2D_s\tilde{\boldsymbol{\mathcal{E}}}^{\dagger}_o\frac{\delta^2\mathcal{K}}{\delta\boldsymbol{E}\delta\dot{\boldsymbol{E}}}D_{\bar{s}}\tilde{\boldsymbol{\mathcal{E}}}_o[\{F_{\bar{s}}, \varepsilon_{\bar{s}}\}^{gc}_{\bar{s}}] \\ &\quad + \left(D_s\tilde{\boldsymbol{\mathcal{E}}}^{\dagger}_o\frac{\delta^2\mathcal{K}}{\delta F_{\bar{s}}\delta\dot{\boldsymbol{E}}} - \frac{\delta^2\mathcal{K}}{\delta F_s\delta\dot{\boldsymbol{E}}}^{\dagger}D_{\bar{s}}\tilde{\boldsymbol{\mathcal{E}}}_o\right)[\{F_{\bar{s}}, \varepsilon_{\bar{s}}\}^{gc}_{\bar{s}}].\end{aligned} \tag{B 2}$$

Here,

$$D_{\bar{s}}\tilde{\boldsymbol{\mathcal{E}}}_o[\delta F_{\bar{s}}] = \frac{1}{\epsilon_o}\boldsymbol{\nabla}G_{\perp}[q_{\bar{s}}\delta N_{\bar{s}}] - \frac{1}{\epsilon_o}\boldsymbol{\nabla}G_{\perp}\left[\boldsymbol{\nabla}\cdot\left(\frac{c^2\mu_o m_{\bar{s}}\delta N_{\bar{s}}}{B^2}\nabla_{\perp}G_{\perp}[\rho(F)]\right)\right] \tag{B 3}$$

$$D_s\tilde{\boldsymbol{\mathcal{E}}}^{\dagger}_o[\boldsymbol{u}] = -\frac{q_s}{\epsilon_o}G_{\perp}[\boldsymbol{\nabla}\cdot\boldsymbol{u}] - \frac{c^2\mu_o m_s}{B^2}\tilde{\boldsymbol{\mathcal{E}}}_o\cdot\nabla_{\perp}G_{\perp}[\boldsymbol{\nabla}\cdot\boldsymbol{u}] \tag{B 4}$$

$$\frac{\delta^2\mathcal{K}}{\delta\boldsymbol{E}\delta\dot{\boldsymbol{E}}}[\delta\boldsymbol{E}] = \sum_s\frac{m_s N_s}{2B^2\omega_{cs}}\hat{b}\times\delta\boldsymbol{E} \tag{B 5}$$

$$\frac{\delta^2\mathcal{K}}{\delta F_{\bar{s}}\delta\dot{\boldsymbol{E}}}[\delta F_{\bar{s}}] = \frac{m_{\bar{s}}}{2B^2\omega_{c\bar{s}}}\hat{b}\times\tilde{\boldsymbol{\mathcal{E}}}_o\delta N_{\bar{s}} \tag{B 6}$$

$$\frac{\delta^2\mathcal{K}}{\delta F_s\delta\dot{\boldsymbol{E}}}^{\dagger}[\delta\boldsymbol{E}] = \frac{m_s}{2B^2\omega_{cs}}\hat{b}\times\tilde{\boldsymbol{\mathcal{E}}}_o\cdot\delta\boldsymbol{E}. \tag{B 7}$$

Consider the problem of inverting the operator $\mathcal{D}$, i.e. solving the equation

$$\sum_{\bar{s}}\mathcal{D}_{s\bar{s}}[\varepsilon_{\bar{s}}] = \chi_s \tag{B 8}$$

for ε given χ. I will argue that this can be done provided $|e|^{-1}$ is small enough and $\rho(F) = \sum_s q_s N_s = O(1)$ (the latter condition being necessary in high-flow drift kinetics in order to guarantee $\boldsymbol{u}_E = O(1)$).

As a first step toward solving for ε, I will solve for the quantity $\langle\delta\varepsilon\rangle$ as a function of $\langle\delta\chi\rangle$, where

$$\langle\delta\varepsilon\rangle_s = \int\{F_s, \varepsilon_s\}^{gc}_s\mathcal{J}_s\mathrm{d}v_{\parallel}\,\mathrm{d}\mu\,\mathrm{d}\zeta \tag{B 9}$$

$$\langle\delta\chi\rangle_s = \int\{F_s, \chi_s\}^{gc}_s\mathcal{J}_s\,\mathrm{d}v_{\parallel}\,\mathrm{d}\mu\,\mathrm{d}\zeta. \tag{B 10}$$

Note that, according to (B 2), (1) the quantity $\eta[\delta F]$ only depends on δF through the velocity-space integral $\int\delta F_s\mathcal{J}_s\,\mathrm{d}v_{\parallel}\,\mathrm{d}\mu\,\mathrm{d}\zeta$ and (2) $\eta[\delta F]$ has no velocity dependence.

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

Therefore, there is a unique operator, $\langle\eta\rangle$, mapping functions on configuration space into itself that satisfies

$$\eta_{s\bar{s}}[\delta F_{\bar{s}}] = \langle\eta\rangle_{s\bar{s}}\left[\int \delta F_{\bar{s}}\mathcal{J}_{\bar{s}}\,\mathrm{d}v_\parallel\,\mathrm{d}\mu\,\mathrm{d}\zeta\right]. \tag{B 11}$$

The equation (B 8) can therefore be expressed as

$$\varepsilon_s - \sum_{\bar{s}}\langle\eta\rangle_{s\bar{s}}[\langle\delta\varepsilon\rangle_{\bar{s}}] = \chi_s, \tag{B 12}$$

which implies

$$\sum_{\bar{s}}\langle\mathcal{D}\rangle_{s\bar{s}}[\langle\delta\varepsilon\rangle_{\bar{s}}] = \langle\delta\chi\rangle_s, \tag{B 13}$$

where

$$\langle\mathcal{D}\rangle_{s\bar{s}}[\langle\delta\varepsilon\rangle_{\bar{s}}] = \delta_{s\bar{s}}\langle\varepsilon\rangle_{\bar{s}} - \int\{F_s,\,\langle\eta\rangle_{s\bar{s}}[\langle\delta\varepsilon\rangle_{\bar{s}}]\}^{gc}_s\mathcal{J}_s\,\mathrm{d}v_\parallel\,\mathrm{d}\mu\,\mathrm{d}\zeta. \tag{B 14}$$

I will now argue that the operator $\langle\mathcal{D}\rangle$ is invertible as a mapping from functions on configuration space into itself. The general principal I will apply asserts that operators sufficiently close to invertible operators are also invertible. The operator $\langle\mathcal{D}\rangle$ can be decomposed as $\langle\mathcal{D}\rangle = \langle\mathcal{D}\rangle_o + \langle\mathcal{D}\rangle_1$ where $\langle\mathcal{D}\rangle_o$ is the leading-order (in $\epsilon = |e|^{-1}$) contribution to $\langle\mathcal{D}\rangle$ and $\langle\mathcal{D}\rangle_1 = O(\epsilon)$. $\langle\mathcal{D}\rangle$ is invertible because $\langle\mathcal{D}\rangle_o$ is invertible and ϵ-close to $\langle\mathcal{D}\rangle$. To see that $\langle\mathcal{D}\rangle_o$ is invertible, first note that

$$\langle\mathcal{D}\rangle_{os\bar{s}}[\langle\delta\varepsilon\rangle_{\bar{s}}] = \delta_{s\bar{s}}\langle\delta\varepsilon\rangle_{\bar{s}} - \boldsymbol{\nabla}\cdot\left(\frac{\hat{b}\times\boldsymbol{\nabla}\langle\eta\rangle_{os\bar{s}}[\langle\delta\varepsilon\rangle_{\bar{s}}]}{q_sB}N_s\right), \tag{B 15}$$

where

$$\begin{aligned}\langle\eta\rangle_{os\bar{s}}[\langle\delta\varepsilon\rangle_{\bar{s}}] &= -2q_s\frac{G_\perp}{\epsilon_o}\left[\boldsymbol{\nabla}\cdot\left(\sum_\sigma\frac{m_\sigma N_\sigma}{2B^2\omega_{c\sigma}}\hat{b}\times\boldsymbol{\nabla}\frac{G_\perp}{\epsilon_o}[q_{\bar{s}}\langle\delta\varepsilon\rangle_{\bar{s}}]\right)\right]\\ &\equiv q_sA[q_{\bar{s}}\langle\delta\varepsilon\rangle_{\bar{s}}],\end{aligned} \tag{B 16}$$

and A is an $O(\epsilon)$ linear operator. Now consider solving the equation,

$$\sum_{\bar{s}}\langle\mathcal{D}\rangle_{os\bar{s}}[\langle\delta\varepsilon\rangle_{\bar{s}}] = j_s, \tag{B 17}$$

for $\langle\delta\varepsilon\rangle$ given j. In terms of the operator A, this equation can be written

$$\langle\delta\varepsilon\rangle_s - \boldsymbol{\nabla}\cdot\left(\hat{b}\times\boldsymbol{\nabla}A\left[\sum_{\bar{s}}q_s\langle\delta\varepsilon\rangle_{\bar{s}}\right]\frac{N_s}{B}\right) = j_s, \tag{B 18}$$

which implies that the quantity $\lambda = \sum_{\bar{s}}q_{\bar{s}}\langle\delta\varepsilon\rangle_{\bar{s}}$ satisfies

$$\lambda + L[\lambda] = \sum_{\bar{s}}q_{\bar{s}}j_{\bar{s}}, \tag{B 19}$$

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

where L is the $O(\epsilon)$ operator given by

$$L[\lambda] = -\nabla \cdot \left(\hat{b} \times \nabla A[\lambda] \frac{\rho(F)}{B} \right). \tag{B 20}$$

Because $L = O(\epsilon)$, the operator $(1 + L)$ is invertible for ϵ small enough. Thus, $\lambda = (1 + L)^{-1}[\sum_{\bar{s}} q_{\bar{s}} j_{\bar{s}}]$, and, according to (B 18),

$$\langle \delta\varepsilon \rangle_s = j_s + \nabla \cdot \left(\hat{b} \times \nabla A \left[(1 + L)^{-1} \left[\sum_{\sigma} q_\sigma j_\sigma \right] \right] \frac{N_s}{B} \right), \tag{B 21}$$

which shows that $\langle \mathcal{D} \rangle_o$ is invertible for small enough ϵ, as claimed. For reasons already mentioned, it follows that $\langle \mathcal{D} \rangle$ is invertible for small enough ϵ. Therefore, by (B 13), $\langle \delta\varepsilon \rangle = \langle \mathcal{D} \rangle^{-1}[\langle \delta\chi \rangle]$. When this result is substituted into (B 12), an expression for ε and a function of χ emerges,

$$\varepsilon = \chi + \langle \eta \rangle \langle \mathcal{D} \rangle^{-1}[\langle \delta\chi \rangle], \tag{B 22}$$

which shows that $\mathcal{D}$ is invertible.

REFERENCES

Abraham, R. & Marsden, J. E. 2008 *Foundations of Mechanics*. AMS Chelsea.

Bourgain, J. 1996 Invariant measures for the 2d-defocusing nonlinear Schrödinger equation. *Commun. Math. Phys.* **176** (2), 421–445.

Brizard, A. J. 2000 Variational principle for nonlinear gyrokinetic Vlasov–Maxwell equations. *Phys. Plasmas* **7**, 4816.

Brizard, A. J. & Hahm, T. S. 2007 Foundations of nonlinear gyrokinetic theory. *Rev. Mod. Phys.* **79**, 421–468.

Burby, J. W. 2015 arXiv:1509.04551.

Burby, J. W. & Brizard, A. J. 2015 Energetically consistent collisional gyrokinetics. *Phys. Plasmas* **22**, 100707.

Burby, J. W., Brizard, A. J., Morrison, P. J. & Qin, H. 2015 Hamiltonian gyrokinetic Vlasov–Maxwell system. *Phys. Lett.* A **379**, 2073–2077.

Cary, J. R. & Littlejohn, R. G. 1983 Noncanonical Hamiltonian mechanics and its application to magnetic field line flow. *Ann. Phys.* **151**, 1–34.

Cendra, H., Holm, D. D., Hoyle, M. J. W. & Marsden, J. E. 1998 The Maxwell–Vlasov equations in Euler–Poincaré form. *J. Math. Phys.* **39**, 3138–3157.

Craig, W. & Wayne, C. E. 1994 Periodic solutions of nonlinear Schrödinger equations and the Nash–Moser method. In *Hamiltonian Mechanics: Integrability and Chaotic Behavior*, p. 103. Springer.

Holm, D. D., Marsden, J. E. & Ratiu, T. S. 1998 The Euler–Poincaré equations and semidirect products with applications to continuum theories. *Adv. Math.* **137**, 1–81.

Holm, D. D. & Tronci, C. 2012 Euler–Poincaré formulation of hybrid plasma models. *Commun. Math. Sci.* **10** (1), 191–222.

Jones, C. K. R. T. 1995 *Geometric Singular Perturbation Theory*, Lecture Notes in Mathematics, vol. 1609, pp. 44–118. Springer.

Krommes, J. A. & Hammett, G. W. 2013 Report of the study group GK2 on momentum transport in gyrokinetics. PPPL Report PPPL-4945, Princeton University.

Kruger, S. E., Hegna, C. C. & Callen, J. D. 1998 Generalized reduced magnetohydrodynamic equations. *Phys. Plasmas* **5**, 4169.

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press

LEE, J. & WALD, R. M. 1990 Local symmetries and constraints. *J. Math. Phys.* **31**, 725.

MARSDEN, J. E., PATRICK, G. W. & SHKOLLER, S. 1998 Multisymplectic geometry, variational integrators, and nonlinear PDEs. *Commun. Math. Phys.* **199**, 351–395.

MIYATO, N., SCOTT, B. D., STRINTZI, D. & TOKUDA, S. 2009 A modification of the guiding-centre fundamental 1-form with strong $E \times B$ flow. *J. Phys. Soc. Japan* **78**, 104501.

MORRISON, P. J. & KOTSCHENREUTHER, M. 1990 *The Energy Principle, Negative Energy Modes, and Stability*, chap. 7, p. 910. World Scientific.

MORRISON, P. J. & PFIRSCH, D. 1989 Free energy expressions for Vlasov equilibria. *Phys. Rev.* A **40**, 3898.

MOTOHASHI, H. & SUYAMA, T. 2015 Third order equations of motion and the Ostrogradsky instability; doi:10.1103/PhysRevD.91.085009.

OSTROGRADSKY, M. V. 1850 *Mem. Acad. St. Petersbourg* **VI 4**, 385.

PARRA, F. I., BARNES, M., CALVO, I. & CATTO, P. J. 2012 Intrinsic rotation with gyrokinetic models. *Phys. Plasmas* **19**, 056116.

PARRA, F. I. & CATTO, P. J. 2008 Limitations of gyrokinetics on transport time scales. *Plasma Phys. Control. Fusion* **50**, 065014.

RAMOS, J. J. 2008 Finite-larmor-radius kinetic theory of a magnetized plasma in the macroscopic flow reference frame. *Phys. Plasmas* **515**, 082106.

SCOTT, B. 2016 Gyrokinetic theory and dynamics of the tokamak edge. *Contrib. Plasma Phys.* (in press); doi:10.1002/ctpp.201010039.

SCOTT, B. & SMIRNOV, J. 2010 Energetic consistency and momentum conservation in the gyrokinetic description of tokamak plasmas. *Phys. Plasmas* **17**, 112302.

SQUIRE, J., QIN, H., TANG, W. M. & CHANDRE, C. 2013 The Hamiltonian structure and Euler–Poincaré formulation of the Vlasov–Maxwell and gyrokinetic systems. *Phys. Plasmas* **20**, 022501.

SUGAMA, H. 2000 Gyrokinetic field theory. *Phys. Plasmas* **7**, 466.

TRONCI, C. 2010 Hamiltonian approach to hybrid plasma models. *J. Phys. A: Math. Theor.* **43**, 375501.

VANNESTE, J. & WIROSOETISNO, D. 2006 Two-dimensional Euler flows in slowly deforming domains. arXiv:math/0611875.

WAYNE, C. E. 1990 Periodic and quasi-periodic solutions of nonlinear wave equations via kam theory. *Commun. Math. Phys.* **127** (3), 479–528.

WENTZEL, G. 1949 *Quantum Theory of Fields*. Interscience.

https://doi.org/10.1017/S0022377816000532 Published online by Cambridge University Press