A NEW OPTIMALITY PROPERTY OF STRANG'S SPLITTING *

FERNANDO CASAS†, JESÚS MARÍA SANZ-SERNA‡, AND LUKE SHAW†

Abstract. For systems of the form $\dot{q} = M^{-1}p$, $\dot{p} = -Aq + f(q)$, common in many applications, we analyze splitting integrators based on the (linear/nonlinear) split systems $\dot{q} = M^{-1}p$, $\dot{p} = -Aq$ and $\dot{q} = 0$, $\dot{p} = f(q)$. We show that the well-known Strang splitting is optimally stable in the sense that, when applied to a relevant model problem, it has a larger stability region than alternative integrators. This generalizes a well-known property of the common Störmer/Verlet/leapfrog algorithm, which of course arises from Strang splitting based on the (kinetic/potential) split systems $\dot{q} = M^{-1}p$, $\dot{p} = 0$ and $\dot{q} = 0$, $\dot{p} = -Aq + f(q)$.

Key words. Strang, Splitting Integrators, Hybrid Monte Carlo, Numerical Stability

MSC codes. 65L07, 65L20

This paper is dedicated to Gilbert Strang.

1. Introduction.

We are concerned with numerical integrators for second-order systems in $\mathbb{R}^d$

$$M\ddot{q} = -Aq + f(q), \tag{1.1}$$

where M and A are constant $d \times d$ matrices (M invertible), or equivalently for first-order systems in $\mathbb{R}^{2d}$

$$\dot{q} = M^{-1}p, \qquad \dot{p} = -Aq + f(q).$$

Our aim is to prove that the Strang splitting integrator [32] based on the (linear/nonlinear) split systems

$$\dot{q} = M^{-1}p, \qquad \dot{p} = -Aq \tag{1.2}$$

and

$$\dot{q} = 0, \qquad \dot{p} = f(q) \tag{1.3}$$

possesses an optimal stability property.

The format (1.1) is a particular instance of the system

$$M\ddot{q} = g(q) \tag{1.4}$$

that appears very frequently in many applications. The best-known integrator for (1.4) is perhaps the Störmer/leapfrog/Verlet algorithm [20]. In its Verlet formulation, the integrator is constructed by applying Strang's splitting to the first-order system

$$\dot{q} = M^{-1}p, \qquad \dot{p} = g(q),$$

with the (kinetic/potential) split systems

$$\dot{q} = M^{-1}p, \qquad \dot{p} = 0, \tag{1.5}$$

and

$$\dot{q} = 0, \qquad \dot{p} = g(q). \tag{1.6}$$

More precisely, let us denote by $\varphi_t^{[D]}$ the solution flow of (1.5), $t \in \mathbb{R}$,

$$\varphi_t^{[D]}(q,p) = (q + tM^{-1}p, p),$$

*Accepted for publication February 14, 2023; published electronically June 6, 2023 in SIAM J. Numer. Anal.

Funding: This work was supported by Ministerio de Ciencia e Innovación (Spain) through project PID2019-104927GB-C21, MCIN/AEI/10.13039/501100011033, ERDF ("A way of making Europe").

†Departament de Matemàtiques and IMAC, Universitat Jaume I, E-12071, Castellón, Spain (casas@uji.es, shaw@uji.es).

‡Departamento de Matemáticas, Universidad Carlos III de Madrid, E-28911, Leganés, Spain (jmsanzserna@gmail.com).

and by $\varphi_t^{[K]}$ the solution flow of (1.6),

$$\varphi_t^{[K]}(q,p) = (q, p + tg(q)),$$

then a timestep of length $h > 0$ of the position Verlet algorithm is given by the map

$$\psi_h^{[pos]} = \varphi_{h/2}^{[D]} \circ \varphi_h^{[K]} \circ \varphi_{h/2}^{[D]}$$

and a step of the velocity Verlet algorithm is defined by the map

$$\psi_h^{[vel]} = \varphi_{h/2}^{[K]} \circ \varphi_h^{[D]} \circ \varphi_{h/2}^{[K]},$$

where the roles of $\varphi^{[D]}$ and $\varphi^{[K]}$ have been swapped. The labels D and K we have used correspond to the words *drift* and *kick*, commonly used in molecular dynamics to refer to $\varphi^{[D]}$ and $\varphi^{[K]}$ respectively [18].

In spite of its simplicity, the Verlet integrator is the method of choice in many applications [24]. One of the advantages of the (position or velocity) Verlet integrator is that it possesses, among a wide class of *explicit* integrators, an *optimal* stability interval [22, 16, 30, 10]. In fact, Verlet strictly maximizes the scaled length of the stability interval, i.e. the quotient Λ/m, where Λ is the length of the stability interval and m the number of evaluations of g per step. In other words, for any explicit competitor integrator using m evaluations per step, there are values of h such that Verlet integrations with steplength h are stable while the (equally costly) integrations of the competitor with steplength mh are unstable. In short, the Verlet algorithm may be operated with longer (scaled) timesteps than any of its explicit competitors; this makes it appealing in applications, including molecular dynamics, where integrations are performed with values of h close to the stability limit because high accuracy is either not required or impossible to achieve due to the complexity of the problem (for instance in cases where g is very expensive to evaluate).

When, in (1.4), g takes the particular form $g(q) = -Aq + f(q)$ as in (1.1), instead of splitting the given system as (1.5)–(1.6), it may be advantageous to split as (1.2)–(1.3) and consider the Strang integrators RKR and KRK

$$\psi_h^{[RKR]} = \varphi_{h/2}^{[R]} \circ \varphi_h^{[K]} \circ \varphi_{h/2}^{[R]} \tag{1.7}$$

and

$$\psi_h^{[KRK]} = \varphi_{h/2}^{[K]} \circ \varphi_h^{[R]} \circ \varphi_{h/2}^{[K]}, \tag{1.8}$$

where $\varphi_t^{[R]}$ and $\varphi_t^{[K]}$ denote respectively the solution flows of the systems (1.2) and (1.3). (Of course, kicks are now based on f rather than on g.) We use the identifier R from *rotation* because in typical applications the matrices M and A are symmetric and positive definite and then the solution map

$$\begin{bmatrix} q \\ p \end{bmatrix} \mapsto \exp\left(t \begin{bmatrix} 0 & M^{-1} \\ -A & 0 \end{bmatrix} \right) \begin{bmatrix} q \\ p \end{bmatrix}$$

of (1.2) describes, after a suitable linear change of variables, d rotations in the (two-dimensional) planes (q^i, p^i), $i = 1, \dots, d$, where q^i and p^i are the scalar components of q and p. The splitting (1.2)–(1.3) is particularly appealing when, in $g(q) = -Aq + f(q)$, $f(q)$ is a small perturbation of $-Aq$: RKR, KRK and other splitting algorithms using sequences of rotations and kicks are exact if the perturbation vanishes. The main contribution of this paper is to show that, as is the case for the velocity and position Verlet integrators, the RKR and KRK integrators (1.7)–(1.8) possess an optimal stability property. Roughly speaking, we show that for a model test problem, for each given steplength, RKR and KRK remain stable for larger perturbations f than any other rotation/kick splitting integrator (see Section 3 for a precise statement).

Motivation. Our interest in problems of the form (1.1) originated when studying integrators for the Hamiltonian Monte Carlo (HMC) method, a sampling technique widely used in statistics and statistical physics [26, 28]. The bulk of the computational effort in HMC is in integrating systems of the form (1.4) where $g(q)$ is the negative gradient of the logarithm of the target probability density function and M is a positive-definite symmetric matrix *chosen by the user*. Therefore devising suitable efficient integrators is of key importance to HMC [8, 10]. In many situations of interest [31], the target density is a perturbation of a Gaussian density and then $g(q) = -Aq + f(q)$ with A the symmetric positive-definite precision matrix of the Gaussian distribution and $f(q)$ a perturbation. As shown in [15], it is then very advantageous to choose $M = A$ and then (1.1) becomes

$$\ddot{q} = -q + \bar{f}(q), \qquad \bar{f}(q) = A^{-1} f(q). \tag{1.9}$$

It is also shown in [15] that to integrate (1.1) or (1.9) the Strang splitting is far more efficient when applied to (1.2)–(1.3) than when applied to the kinetic/potential (1.5)–(1.6). This suggests the investigation of rotation/kick splitting algorithms for (1.1) or (1.9). Furthermore, for reasons detailed in [5, 4], as a rule, integrations of (1.9) within HMC simulations are best carried out with values of h close to the stability limit of the integrator. Therefore it is of clear interest to identify the rotation/kick splitting integrators with optimal stability interval. In fact the motivation for the present research originated when our multiple attempts to construct integrators that improved on KRK or RKR failed [15].

Exponential integrators [21] are a well-known class of algorithms that, as splitting methods, exploit the structure of (1.1) or (1.9). However they are not relevant to HMC applications where symplecticness and time-reversibility are essential [10].

Contents. The article has five sections. Section 2 contains preliminary material. The main optimality result is presented and proved in Section 3. Section 4 provides complementary results to compare the size of the stability regions of the Strang splitting algorithms and some possible competitors. The final section contains a technical proposition.

2. Preliminaries.

In this section we present a number of facts that are required to formulate and prove the main result presented in the next section.

2.1. Splitting integrators.

The importance of splitting integrators in different applications has increased substantially in recent decades [6], often in connection with preservation of geometric properties, such as symplecticness [29]. Of course, the RKR and KRK methods (1.7) and (1.8) are not the only splitting algorithms to integrate (1.1) with the help of the split systems (1.2) and (1.3). One may consider m-stage integrators by interleaving rotations and kicks, beginning with either R or K as follows

$$\psi_h = \varphi^{[R]}_{r_{m+1}h} \circ \varphi^{[K]}_{k_m h} \circ \varphi^{[R]}_{r_m h} \circ \dots \circ \varphi^{[K]}_{k_1 h} \circ \varphi^{[R]}_{r_1 h}, \quad \psi_h = \varphi^{[K]}_{k_{m+1}h} \circ \varphi^{[R]}_{r_m h} \circ \varphi^{[K]}_{k_m h} \circ \dots \circ \varphi^{[R]}_{r_1 h} \circ \varphi^{[K]}_{k_1 h}. \tag{2.1}$$

We always assume the consistency requirements $\sum_i r_i = 1$ and $\sum_i k_i = 1$. Some of the coefficients r_i or k_i are allowed to vanish as this simplifies the presentation. Note that the first format in (2.1) uses (at most) m kicks and therefore (at most) $\leq m$ evaluations of f per step; the second format uses $\leq m+1$ kicks, but, since, if $k_{m+1} \neq 0$ and $k_1 \neq 0$, the value of f at the last kick of the current timestep may be used to perform the first kick of the next timestep, also requires essentially $\leq m$ evaluations of f per timestep.

If M and A are symmetric and positive definite and $f(q) = -\nabla V(q)$ for a suitable scalar function V, then (1.1) is equivalent to the Hamiltonian system with Hamiltonian function $(1/2)p^T M^{-1} p + (1/2) q^T A q + V(q)$. In this case the split systems (1.2) and (1.3) are also Hamiltonian and therefore $\varphi^{[R]}_t$ and $\varphi^{[K]}_t$ are, for each $t \in \mathbb{R}$, symplectic maps, as flows of Hamiltonian systems. It follows that the splitting integrators in (2.1) will be symplectic, as is required in HMC applications [10].

It is often the case that the coefficients r_i, k_i in (2.1) are chosen *palindromically*, i.e. for compositions starting with R, $r_{m+2-i} = r_i$, $i = 1, \dots, m+1$, and $k_{m+1-j} = k_j$, $j = 1, \dots, m$, and similarly for compositions starting with K. RKR and KRK are both palindromic. Palindromic splitting integrators have at least second order of accuracy and, in addition, are time-reversible, as required in HMC applications [10].

2.2. Conjugate integrators.

Given two integrators ψ_h and $\bar{\psi}_h$ of the form (2.1), we say that they are conjugate if there is an invertible map χ_h such that

$$\bar{\psi}_h = \chi_h \circ \psi_h \circ \chi_h^{-1}.$$

This notion goes back to Butcher's algebraic theory of Runge-Kutta methods [11, 12, 13]. The n-fold composition map $\bar{\psi}_h^n$ used to advance n steps with method $\bar{\psi}_h$ may be written as

$$\bar{\psi}_h^n = (\chi_h \circ \psi_h \circ \chi_h^{-1}) \circ (\chi_h \circ \psi_h \circ \chi_h^{-1}) \circ \dots \circ (\chi_h \circ \psi_h \circ \chi_h^{-1}) = \chi_h \circ \psi_h^n \circ \chi_h^{-1},$$

and therefore to advance n steps with method $\bar{\psi}_h$ one may (i) apply once the map χ_h^{-1} (preprocessing), (ii) advance n steps with the integrator ψ_h, (iii) apply once the map χ_h (postprocessing). Butcher was interested in the case where $\bar{\psi}_h$ has order of consistency higher than ψ_h, since then pre/postprocessing make it possible to perform high-order integrations with $\bar{\psi}_h$ by implementing the low-order integrator ψ_h.

An example of conjugate methods is afforded by the integrators RKR and KRK with the postprocessor $\chi_h = \varphi_{h/2}^{[R]} \circ \varphi_{h/2}^{[K]}$:

$$
\begin{aligned}
\psi_h^{[RKR]} &= \varphi_{h/2}^{[R]} \circ \varphi_h^{[K]} \circ \varphi_{h/2}^{[R]} \\
&= \left(\varphi_{h/2}^{[R]} \circ \varphi_{h/2}^{[K]}\right) \circ \left(\varphi_{h/2}^{[K]} \circ \varphi_h^{[R]} \circ \varphi_{h/2}^{[K]}\right) \circ \left(\varphi_{h/2}^{[R]} \circ \varphi_{h/2}^{[K]}\right)^{-1} \\
&= \chi_h \circ \psi_h^{[KRK]} \circ \chi_h^{-1}.
\end{aligned}
$$

One may prove by means of similar manipulations that all (consistent) one-stage integrators, including the non palindromic, first-order Lie-Trotter integrators $\varphi_h^{[R]} \circ \varphi_h^{[K]}$ and $\varphi_h^{[K]} \circ \varphi_h^{[R]}$ may be conjugated to either RKR or KRK, which are palindromic and second-order. Clearly, $\varphi_h^{[R]} \circ \varphi_h^{[K]}$ is obtained by setting $r_2 = 1$, $k_1 = 1$, $r_1 = 0$ in the first equality in (2.1); $\varphi_h^{[K]} \circ \varphi_h^{[R]}$ results from the choice $r_2 = 0$, $k_1 = 1$, $r_1 = 1$ in the same equality. Both integrators may also be obtained by using the format in the second equality in (2.1).

It is proved in [7] that every integrator may be conjugated to a palindromic integrator.

For each problem (1.1) the numerical trajectory $\psi_h^n(q,p)$, $n = 0, 1, 2, \ldots$, generated by ψ_h with initial condition (q,p) is mapped by χ_h into the trajectory $\bar{\psi}_h^n(q^*,p^*)$, $n = 0, 1, 2, \ldots$, with initial condition $(q^*,p^*) = \chi_h(q,p)$. For this reason the long-time properties of the numerical solutions generated by ψ_h and $\bar{\psi}_h$ may be expected to be similar (for instance bounded/unbounded trajectories of ψ_h correspond to bounded/unbounded trajectories of $\bar{\psi}_h$).

2.3. The model problem. Roughly speaking, a numerical integration with a given integrator and steplength h is said to be unstable if the numerical solution shows unphysical growth as the number of computed timesteps increases. In order to make this notion mathematically precise, it is standard to restrict the attention to integrations performed on an easy-to-analyse *model* problem chosen in such a way that conclusions based on the model are relevant when dealing with more general problems.

For (1.4), it is standard to use the model scalar problem $\ddot{q} = -\omega^2 q$, i.e. the familiar harmonic oscillator. The relevance of this choice of model problem may be justified as follows. Let us assume, for simplicity, that M, as is the case in most applications, is symmetric and positive-definite (this hypothesis may be relaxed). Writing $M = LL^T$ and introducing new variables $\bar{q} = L^T q$, (1.4) becomes $\ddot{\bar{q}} = L^{-1} g(L^{-T}\bar{q})$. Furthermore, if g is linear, $g(q) = -Aq$, then $\ddot{\bar{q}} = -L^{-1}AL^{-T}\bar{q}$. The important case, with oscillatory solutions, is that where $L^{-1}AL^{-T}$ is diagonalizable with positive eigenvalues (which happens if in particular A is symmetric and positive definite). Then a new change of variables reduces the system to a set of d uncoupled scalar harmonic oscillators $\ddot{q} = -\omega^2 q$ (the eigenvalues of $L^{-1}AL^{-T}$ provide the values of ω^2). For this construction to be useful it is required that the transformations that diagonalize the system being integrated also diagonalize the integrator, something that invariably happens for all integrators of practical interest.

In order to identify a suitable model problem for integrators for (1.1) we proceed similarly. We consider the case where f is linear $f(q) = -Bq$; the change of variables $\bar{q} = L^T q$ brings the system to the form $\ddot{\bar{q}} = -L^{-1}(A+B)L^{-T}\bar{q}$. Under the hypothesis that there is a linear transformation that brings both $L^{-1}AL^{-T}$ and $L^{-1}BL^{-T}$ to diagonal form, after a new change of variables the system is transformed into d uncoupled scalar equations of the form

$$\ddot{q} = -(\lambda + \mu) q, \tag{2.2}$$

where λ and μ are eigenvalues of $L^{-1}AL^{-T}$ and $L^{-1}BL^{-T}$ associated with the same eigenvector. We are interested in problems with $\lambda > 0$ and $\lambda + \mu > 0$ (something which happens in the important case where A and $A + B$ are symmetric and positive definite), so that the equations (2.2) corresponds to harmonic oscillators. The analysis of (2.2) is simplified if we introduce a new time variable $t/\sqrt{\lambda}$, so as to have, after denoting $\varepsilon = \mu/\lambda$,

$$\ddot{q} = -q - \varepsilon q, \qquad \varepsilon > -1. \tag{2.3}$$

This model problem, that we refer to hereafter as "the model problem", has appeared e.g. in [9].

In the particular situation of the system (1.9) arising in the HMC method, the derivation just outlined of the model (2.3) may be greatly simplified. In fact, if f is linear, $f(u) = -Bu$ so that $\bar{f}(u) = A^{-1}Bu$, and $A^{-1}B$ diagonalizes with eigenvalues $\varepsilon > -1$, then a single change of variables reduces (1.9) to d uncoupled harmonic oscillators of the form (2.3). In the case where $f(u) = -Bu$ is a small perturbation of Au, the eigenvalues ε will actually have small magnitude.

2.4. Integrating the model problem. Stability. For the model problem (2.3),

$$\varphi_t^{[R]}(q,p) = \begin{bmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{bmatrix} \begin{bmatrix} q \\ p \end{bmatrix}, \qquad \varphi_t^{[K]}(q,p) = \begin{bmatrix} 1 & 0 \\ -t\varepsilon & 1 \end{bmatrix} \begin{bmatrix} q \\ p \end{bmatrix},$$

where we note that both transformations have unit determinant as each corresponds to the flow of a Hamiltonian system. By multiplying the matrices that represent the flows being composed in (2.1), we obtain the matrices representing one step of the splitting integrator ψ_h. In particular for the Strang splittings (1.7) and (1.8), we find that the matrices that perform a timestep of length h are

$$\begin{bmatrix} \cos(h) - \frac{h\varepsilon}{2}\sin(h) & \sin(h) - \varepsilon h \sin^2\left(\frac{h}{2}\right) \\ -\sin(h) - \varepsilon h \cos^2\left(\frac{h}{2}\right) & \cos(h) - \frac{h\varepsilon}{2}\sin(h) \end{bmatrix} \qquad \text{for} \qquad \psi_{\varepsilon,h}^{[RKR]} \tag{2.4}$$

and

$$\begin{bmatrix} \cos(h) - \frac{h\varepsilon}{2}\sin(h) & \sin(h) \\ -\varepsilon h\cos(h) - (1 - \left(\frac{h\varepsilon}{2}\right)^2)\sin(h) & \cos(h) - \frac{h\varepsilon}{2}\sin(h) \end{bmatrix} \qquad \text{for} \qquad \psi_{\varepsilon,h}^{[KRK]}. \tag{2.5}$$

For the integrators in (2.1) the (real) matrix takes the form

$$M_{\varepsilon,h} = \begin{bmatrix} A_{\varepsilon,h} & B_{\varepsilon,h} \\ C_{\varepsilon,h} & D_{\varepsilon,h} \end{bmatrix}.$$

The dependence of the coefficients $A - D$ on ε is *polynomial* and with m stages A and D are polynomials of degree $\leq m$ in ε (this is easily proved by induction). The dependence on h, on the other hand, involves both powers of h and trigonometric functions, as illustrated by (2.4) and (2.5). For palindromic compositions $A_{\varepsilon,h} = D_{\varepsilon,h}$ (see e.g. [8, 14]).

The matrix $M_{\varepsilon,h}$ has unit determinant, as it results from multiplying rotations and kicks of unit determinant. Then its (possibly complex) eigenvalues are inverse to one another, $\lambda_{\varepsilon,h}$ and $1/\lambda_{\varepsilon,h}$, and it is well known that one of the three following situations obtains:

1. The modulus of the trace $A_{\varepsilon,h} + D_{\varepsilon,h} = \lambda_{\varepsilon,h} + 1/\lambda_{\varepsilon,h}$ of $M_{\varepsilon,h}$ is < 2. This corresponds to two different complex eigenvalues of unit modulus. As n increases the powers $M_{\varepsilon,h}^n$ remain bounded and the integration is *stable*.
2. The modulus of the trace is $= 2$. Then there is a double real eigenvalue $\lambda = 1/\lambda \in \{-1, 1\}$. If, in addition $M_{\varepsilon,h}$ diagonalizes, then $M_{\varepsilon,h}$ is either I (the identity matrix) or $-I$, with bounded powers, and the integration is *stable*. When $M_{\varepsilon,h}$ does not diagonalize its powers grow linearly and the integration is *linearly unstable*.
3. The modulus of the trace is > 2. Then there is one real eigenvalue of modulus > 1, leading to *exponential instability*.

Cases 1 and 3 above are robust against perturbations, in the sense that if, for a given integrator, the pair (ε, h) is in case 1 (respectively, case 3), all sufficiently close pairs are also in case 1 (respectively, case 3). Perturbations of case 2, on the contrary, will generically lead to either case 1 or case 3. The stability region of an integrator is the set in the (ε, h) plane where it is stable.

The semitrace

$$P(\varepsilon, h) = (1/2)(A_{\varepsilon,h} + D_{\varepsilon,h}) = (1/2)(\lambda_{\varepsilon,h} + 1/\lambda_{\varepsilon,h})$$

of $M_{\varepsilon,h}$ will be called, using a not very precise terminology, the *stability polynomial* of the integrator; recall that it is a polynomial in ε of degree $\leq m$ but its dependence on h includes trigonometric functions. Exponentially unstable integrations correspond then to $|P(\varepsilon, h)| > 1$.

If the integrators ψ_h and $\bar{\psi}_h$ are conjugate to each other, then the corresponding matrices satisfy the similarity condition

$$\bar{M}_{\varepsilon,h} = S_{\varepsilon,h} M_{\varepsilon,h} S_{\varepsilon,h}^{-1}$$

where the matrix $S_{\varepsilon,h}$ corresponds to the postprocessor. As a consequence $\bar{M}_{\varepsilon,h}$ and $M_{\varepsilon,h}$ share the same pair of eigenvalues $\lambda_{\varepsilon,h}$, $1/\lambda_{\varepsilon,h}$ and therefore *conjugate integrators share a common stability polynomial.* This property is illustrated by the RKR, KRK pair in (2.4)–(2.5). The property was perhaps to be expected, because it was pointed out above that for any two conjugate integrators the numerical trajectories of one of them are mapped by the processor into numerical trajectories of the other.

2.5. A property of the stability polynomial. The following result will be essential to prove our main result.

PROPOSITION 2.1. *For each (consistent) integrator* (2.1) *the stability polynomial satisfies:*

$$P(\varepsilon, h) = \frac{1}{2}(A_{\varepsilon,h} + D_{\varepsilon,h}) = \cos(h) - \frac{\varepsilon h}{2}\sin(h) + \mathcal{O}(\varepsilon^2), \qquad \varepsilon \to 0. \tag{2.6}$$

Proof. It is sufficient to consider the R-first format in the first equality in (2.1); a K-first integrator may be rewritten in the R-first format by adding dummy rotations of duration $0h$ at the beginning and end of the step. We introduce the matrices

$$R = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \qquad K = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix},$$

whose exponentials represent the rotation and the kick

$$\exp(tR) = \begin{bmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{bmatrix}, \qquad \exp(tK) = I + tK = \begin{bmatrix} 1 & 0 \\ -t & 1 \end{bmatrix}.$$

Then the matrix associated with the integrator is

$$M_{\varepsilon,h} = \exp(hr_{m+1}R)(I + \varepsilon h k_m K)\exp(hr_m R)(I + \varepsilon h k_{m-1} K)\cdots(I + \varepsilon h k_1 K)\exp(hr_1 R), \tag{2.7}$$

which leads to

$$M_{\varepsilon,h} = \exp(h\theta_{m+1}R) + \varepsilon h \sum_{i=1}^{m} k_i \exp(h(1-\theta_i)R) K \exp(h\theta_i R) + \mathcal{O}(\varepsilon^2),$$

where $\theta_i = \sum_{j=1}^{i} r_j$. By consistency $\theta_{m+1} = 1$ and therefore the semitrace of $\exp(h\theta_{m+1}R)$ is $\cos(h)$; this gives the term independent of ε in the stability polynomial, as it was to be established in order to prove (2.6). The term of first degree in ε in the last display may be computed as

$$-\varepsilon h \sum_{i=1}^{m} k_i \begin{bmatrix} \sin(h\theta_i)\cos(h(1-\theta_i)) & \sin(h\theta_i)\sin(h(1-\theta_i)) \\ \cos(h\theta_i)\cos(h(1-\theta_i)) & \cos(h\theta_i)\sin(h(1-\theta_i)) \end{bmatrix}.$$

Thus the coefficient of ε in the stability polynomial is

$$-\frac{h}{2}\sum_{i=1}^{m} k_i \Big(\sin(h\theta_i)\cos(h(1-\theta_i)) + \cos(h\theta_i)\sin(h(1-\theta_i)) \Big) = -\frac{h}{2}\sum_{i=1}^{m} k_i \sin(h) = -\frac{h}{2}\sin(h),$$

as was to be proved. ∎

2.6. Stability of the integrators RKR and KRK. We now study the stability of RKR/KRK with stability polynomial/semitrace (see (2.4)–(2.5)):

$$P(\varepsilon, h) = \cos(h) - \frac{h\varepsilon}{2}\sin(h). \tag{2.8}$$

The conditions $P(\varepsilon,h)=1$ and $P(\varepsilon,h)=-1$ correspond to $\varepsilon=\alpha(h)$ and $\varepsilon=\beta(h)$ respectively with

$$\alpha(h) = -\frac{2}{h}\tan\left(\frac{h}{2}\right), \qquad \beta(h) = \frac{2}{h}\cot\left(\frac{h}{2}\right). \tag{2.9}$$

If we restrict attention to $0<h<\pi$, then the condition $|P(\varepsilon,h)|\leq 1$ holds if and only if $\varepsilon\in[\alpha(h),\beta(h)]$; also, for such values of h, $\alpha(h)<-1$, $0<\beta(h)$. When integrating the model problem (where $\varepsilon>-1$) we have stability for $\varepsilon\in(-1,\beta(h))$ and exponential instability for $\varepsilon>\beta(h)$. The case $\varepsilon=\beta(h)$ yields linear instability. The function $\beta(h)$ decreases monotonically for $h\in(0,\pi)$ and therefore increasing h results in a decrease of the interval $(0,\beta(h))$ of positive values of ε leading to a stable integration. As $h\uparrow\pi$, we have $\beta(h)\downarrow 0$, the interval $(0,\beta(h))$ approaches the empty set and thus there is little interest in considering $h\geq\pi$ when dealing with RKR and KRK. This coincides with the analysis in [24, §4.2.1], where it is shown that $h=\pi$ is unstable for any non-zero ε.

Since, as pointed out before, all (consistent) one-stage integrators are conjugate to RKR or KRK the discussion above also applies to them. In particular, their stability polynomial is also given by (2.8) (a conclusion that may be reached alternatively from Proposition 2.1, after taking into account that for $m=1$ the stability polynomial is of first degree in ε, so that the term $\mathcal{O}(\varepsilon^2)$ in (2.6) must vanish).

2.7. The RKRm and KRKm integrators. To avoid duplications, the presentation in this subsection is limited to RKR, but all the results apply to KRK in an obvious manner.

In the analysis in the next section we shall use the auxiliary m-stage integrator

$$\psi_h^{[RKRm]} = \left(\psi_{h/m}^{[RKR]}\right)^m;$$

a single step of length h of $\psi^{[RKRm]}$ demands performing m consecutive substeps with $\psi^{[RKR]}$, each of steplength h/m. As a consequence, integrations with $\psi^{[RKRm]}$ are in fact nothing but $\psi^{[RKR]}$ integrations; $\psi^{[RKRm]}$ is just a mathematical construction to facilitate the fair comparison between m-stage integrators (with m evaluations of f per step) and the one-stage $\psi^{[RKR]}$ (with only one evaluation of f per step).

Clearly

$$M_{\varepsilon,h}^{[RKRm]} = \left(M_{\varepsilon,h/m}^{[RKR]}\right)^m,$$

and, for the eigenvalues, $\lambda_{\varepsilon,h}^{[RKRm]} = \left(\lambda_{\varepsilon,h/m}^{[RKR]}\right)^m$. It follows easily from (2.9) that, restricting attention to $h<m\pi$, $|P^{[RKRm]}|<1$ if and only if $\varepsilon\in(\alpha_m(h),\beta_m(h))$ with

$$\alpha_m(h) = -\frac{2m}{h}\tan\left(\frac{h}{2m}\right) < -1, \qquad \beta_m(h) = \frac{2m}{h}\cot\left(\frac{h}{2m}\right) > 0. \tag{2.10}$$

When integrating the model problem, RKRm is stable if and only if $\varepsilon\in(\alpha_m(h),\beta_m(h))$ (although, as mentioned above, only stability for $\varepsilon>-1>\alpha_m(h)$ is significant). The case $\varepsilon>\beta_m(h)$ yields exponential instability and $\varepsilon=\beta_m(h)$ gives linear instability. See Figure 1.

We now find an expression for the stability polynomial $P^{[RKRm]}(\varepsilon,h)$. Write $\lambda_{\varepsilon,h}^{[RKR]}=\exp(\mathrm{i}\theta_{\varepsilon,h})$ (θ is real if λ has unit modulus) with i the imaginary unit. Then, recalling (2.8), we may write

$$\cos(h) - \frac{h\varepsilon}{2}\sin(h) = P^{[RKR]}(\varepsilon,h) = \frac{1}{2}\left(\lambda_{\varepsilon,h}^{[RKR]} + \frac{1}{\lambda_{\varepsilon,h}^{[RKR]}}\right) = \frac{1}{2}\big(\exp(\mathrm{i}\theta_{\varepsilon,h}) + \exp(-\mathrm{i}\theta_{\varepsilon,h})\big) = \cos(\theta_{\varepsilon,h}),$$

and

$$P^{[RKRm]}(\varepsilon,h) = \frac{1}{2}\left(\lambda_{\varepsilon,h}^{[RKRm]} + \frac{1}{\lambda_{\varepsilon,h}^{[RKRm]}}\right) = \frac{1}{2}\big(\exp(\mathrm{i}m\theta_{\varepsilon,h/m}) + \exp(-\mathrm{i}m\theta_{\varepsilon,h/m})\big) = \cos(m\theta_{\varepsilon,h/m}),$$

so that, introducing the standard Chebyshev polynomial of the first kind T_m with $T_m(\cos(\zeta))=\cos(m\zeta)$ for all (real or complex) ζ, we conclude that

$$P^{[RKRm]}(\varepsilon,h) = T_m\left(\cos\left(\frac{h}{m}\right) - \frac{h\varepsilon}{2m}\sin\left(\frac{h}{m}\right)\right). \tag{2.11}$$

m	1	2	3	4	5	6	7	8	9	10
h_m	π	4.92	5.98	6.85	7.61	8.30	8.93	9.53	10.08	10.61

Table 1: Values of the quantity h_m used in the main theorem.

3. Main result. In the statement of the main result we denote by h_m the smallest positive root of the equation

$$\frac{mh}{2}\sin\left(\frac{h}{m}\right)=\cos\left(\frac{\pi}{m}\right)-\cos\left(\frac{h}{m}\right).$$

For $m=1$, $h_m=\pi$ and, for $m>1$, $h_m<m\pi$. In addition h_m increases monotonically with m and a straightforward Taylor expansion shows that, as $m\uparrow\infty$, $h_m=12^{1/4}\pi^{1/2}m^{1/2}+o(m^{1/2})$. See Table 1.

Theorem 3.1. *Define h_m as above. Then:*

- *For $h<m\pi$, $\varepsilon>-1$, integrations of the model problem* (2.3) *with either RKRm and KRKm are exponentially unstable if and only if $\varepsilon\in(\beta_m(h),\infty)$.*
- *Consider an m-stage splitting integrator ψ_h of the form* (2.1) *with stability polynomial different from the stability polynomial* (2.11) *of the integrators RKRm/KRKm. Then, for $h\neq\pi,2\pi,\dots,(m-1)\pi$ and $h<h_m$, the (open) set of values of $\varepsilon>-1$ that lead to exponentially unstable integrations of the model problem is* strictly larger *than the interval $(\beta_m(h),\infty)$ where RKRm and KRKm show exponential instability.*

This result may be restated by saying that for each fixed h^*, $h^*<h_m$, $h^*\neq\pi,2\pi,\dots,(m-1)\pi$, the intersection of the stability region with the line $h=h^*$ is strictly larger for RKRm and KRKm than for integrators with stability polynomial different from (2.11). Before we prove Theorem 3.1, we need an auxiliary result that we present in the following subsection.

3.1. Chebyshev polynomials. It is well known that many properties of the Chebyshev polynomials are a consequence of the following equioscillation property: $T_m(\xi_i)=(-1)^i$ at the points $\xi_i=\cos(i\pi/m)$, $i=0,\dots,m$, that partition $[-1,1]$ as $-1=\xi_m<\xi_{m-1}<\cdots<\xi_1<\xi_0=1$. We shall need the following well-known, elementary equioscillation result, whose proof we provide for completeness:

Lemma 3.2. *Consider $k+1$ real points $x_0<x_1<\cdots<x_k$. If Q is a real polynomial such that either*

$$Q(x_i)\geq 0,\ i\text{ even}\quad\text{and}\quad Q(x_i)\leq 0,\ i\text{ odd},$$

or

$$Q(x_i)\leq 0,\ i\text{ even}\quad\text{and}\quad Q(x_i)\geq 0,\ i\text{ odd},$$

then $Q(x)$ has $\geq k$ zeros (counting multiplicities) in the interval $[x_0,x_k]$.

Proof. Consider the k disjoint intervals

$$J_1=[x_0,x_1),\ J_2=[x_1,x_2),\dots,J_{k-1}=[x_{k-2},x_{k-1}),\ J_k=[x_{k-1},x_k],$$

that partition $[x_0,x_k]$. We first point out that $Q(x)$ must have at least a zero in the closed interval J_k (otherwise $Q(x)$ would be strictly >0 or strictly <0 for $x\in[x_{k-1},x_k]$, in contradiction with the hypothesis). On the other hand, it is possible that some of the semiclosed J_i, $i=1,\dots,k-1$, contain no zero of $Q(x)$, but, if that is the case, then $Q(x_i)=0$. Furthermore, in that case, J_{i+1} must contain at least two zeros, for if it only contained a single zero at x_i, then either $Q(x_{i-1})>0$, $Q(x_{i+1})<0$ or $Q(x_{i-1})<0$, $Q(x_{i+1})>0$. Thus, if a subinterval other than J_k carries no zero, then the one to its right carries two, and this gives a total of at least k zeros. □

The following result on Chebyshev polynomials is to our best knowledge not available in the literature. Its proof is based on the preceding lemma.

Proposition 3.3. *For given $m\geq 2$, let $P(x)$ be a real polynomial of degree $\leq m$ different from $T_m(x)$. Assume that $P(x)-T_m(x)$ has a double zero $\xi\in(-1,1)$ such that $\xi\neq\xi_i=\cos(i\pi/m)$ for $i=1,\dots,m-1$. Then $|P(x)|>1$ for some $x\in(\xi_m,\xi_1)=(-1,\cos(\pi/m))$.*

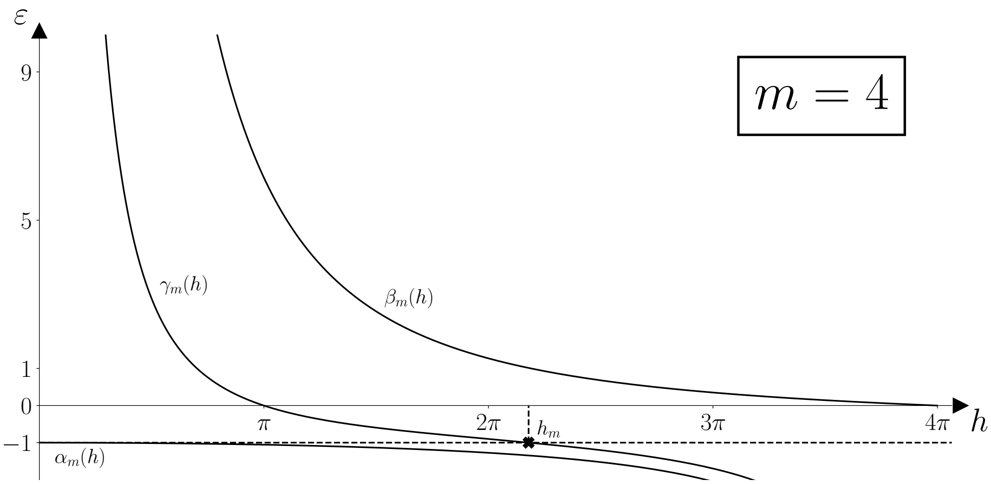

Fig. 1: Proof of the main result in the case $m = 4$. In the model problem $\varepsilon > -1$. RKR4 and KRK4 are stable in the open region bounded by the lines $h = 0$, $h = m\pi$, $\varepsilon = -1$, $\varepsilon = \beta_m(h)$. For each fixed h such that $h < 4\pi$, $h \neq \pi, 2\pi, 3\pi$, a competitor integrator will have $|P(\varepsilon, h)| > 1$ for some $\varepsilon \in (\gamma_h, \beta_h)$. When $h < h_m$, those values of ε are > -1.

Proof. Assume that $|P(x)| \leq 1$ in (ξ_m, ξ_1) and consider the difference $D(x) = P(x) - T_m(x)$. For $i = 1, \ldots, m$ with i odd, we have $D(\xi_i) = P(\xi_i) - T(x_i) = P(\xi_i) - (-1) \geq -1 + 1 = 0$. Similarly, for $i = 1, \ldots, m$ with i even, we have $D(\xi_i) \leq 0$. There are two cases:

1. $\xi \in (\xi_1, \xi_0)$. Then, by the lemma, $D(x)$ has $\geq m - 1$ zeros in $[\xi_m, \xi_1]$. These and the double zero $\xi \in (\xi_1, \xi_0)$ provide $\geq m+1$ zeros of $D(x)$. It follows that $D(x)$ vanishes identically, in contradiction with the hypotheses of the proposition.
2. ξ is in an interval (ξ_{j+1}, ξ_j) with $j = 1, \ldots, m-1$. By applying the lemma twice, we see that $D(x)$ has $\geq j - 1$ zeros in $[\xi_j, \xi_1]$ and $\geq m - j - 1$ zeros in $[\xi_m, \xi_{j+1}]$. The subinterval (ξ_{j+1}, ξ_j) must contain at least three zeros, because, if the multiplicity of ξ were exactly 2 and there were no other zeros in the subinterval, then $D(\xi_j)$ and $D(\xi_{j+1})$ would be either both > 0 or both < 0. We have thus found $\geq j - 1 + (m - j - 1) + 3 = m + 1$ zeros, which again leads to a contradiction. ∎

3.2. Proof of the main result. The first item in Theorem 3.1 was established at the very end of Section 2. In the second item, we only have to deal with $m \geq 2$, because we also saw in Section 2 that there is no consistent one-stage integrator with stability polynomial different from the stability polynomial (2.8) of RKR or KRK.

With fixed h satisfying the conditions of the theorem, we change variables replacing ε by the new variable

$$x = \cos\left(\frac{h}{m}\right) - \frac{\varepsilon h}{2m} \sin\left(\frac{h}{m}\right).$$

Since $h < h_m < m\pi$, this transformation is bijective. It maps $\varepsilon = \alpha_m(h)$ (see (2.10)) into $x = 1$ and $\varepsilon = \beta_m(h)$ into $x = -1$. The change of variables is chosen in such a way that, according to (2.11), the stability polynomial of RKRm or KRKm is transformed into the Chebyshev polynomial $T_m(x)$.

Denote by $P(x)$ the m-degree polynomial in the variable x resulting from changing variables in the stability polynomial $P(\varepsilon, h)$ of the integrator ψ_h (note that the dependence of $P(x)$ on h has been left out of the notation). By hypothesis, $P(x)$ cannot coincide with $T_m(x)$. From Proposition 2.1, $P(\varepsilon, h) - P^{[RKRm]}(\varepsilon, h)$ as a polynomial in ε has a double root at $\varepsilon = 0$ and accordingly $P(x) - T_m(x)$ has a double zero at the corresponding value of x given by $\xi = \cos(h/m)$. Since h is assumed to be $\neq \pi, 2\pi, \ldots, (m-1)\pi$, ξ is not one of the extrema $\xi_i = \cos(i\pi/m)$, $i = 1, \ldots, m-1$, of $T_m(x)$. Proposition 3.3 reveals that $|P(x)|$ has to exceed 1 as some point $x \in (-1, \cos(\pi/m))$; the corresponding ε-value will be in the interval $(\gamma_m(h), \beta_m(h))$

with

$$\gamma_m(h) = \frac{2m}{h\sin(h/m)}\Big(\cos(h/m) - \cos(\pi/m)\Big).$$

The condition $h < h_m$ implies $\gamma_m(h) > -1$ (see Figure 1). We have thus found values of $\varepsilon \in (-1, \beta_m(h))$ that lead to instability and the proof is complete.

4. Assessing the size of the stability region. The result we have just presented does not provide quantitative information on the size of stability regions in the full (ε, h) plane of the different integrators. In this section, we present a more quantitative analysis; it turns out that Strang integrators have much larger stability regions than their competitors.

4.1. Stability near $\varepsilon = 0$, $h = n\pi$. When $\varepsilon = 0$ all splitting integrators (2.1) are exact and therefore $M_{0,h}$ is the matrix corresponding to a rotation by h radians, with semitrace $P(0, h) = \cos(h)$. If $h > 0$ is not an integer multiple of π, the magnitude of the trace is < 2 and the matrix $M_{0,h}$ is strongly stable [3, sections 25 and 42] and [23] (see also [9]). Accordingly, the integrator is stable in a neighborhood of $(0, h)$. On the other hand, $P(0, n\pi) = (-1)^n$, $n = 1, 2, \ldots$, and perturbations of the parameter values $\varepsilon = 0$, $h = n\pi$ may render the integrator exponentially unstable. For instance, RKRm and KRKm are stable, as we know, in the neighbourhood of $(0, \pi)$, …, $(0, (m-1)\pi)$ but not in the neighbourhood of $(0, m\pi)$ (see Figure 1). We now investigate the stability of general integrators (2.1) in the neighbourhood of the points $(0, n\pi)$, $n = 1, 2, \ldots$

We assume that n is *odd* (the case n even is entirely parallel). Then $P(0, n\pi) = -1$ and a necessary condition for the method to be stable in a neighbourhood of $(0, n\pi)$ is that this point be a minimum of P. Since, for $\varepsilon = 0$, $P(0, h) = \cos(h)$, we have $(\partial/\partial h)P(0, h) = -\sin(h)$ and $(\partial/\partial h)P(0, n\pi) = 0$. In addition, from Proposition 2.1, $(\partial/\partial\varepsilon)P(0, h) = -(h/2)\sin(h)$, and, therefore $(\partial/\partial\varepsilon)P(0, n\pi)) = 0$; we conclude that all integrators satisfy the first-order necessary conditions for $(0, n\pi)$ to be a minimum of P. Turning now to the second-order necessary conditions, from $(\partial^2/\partial h^2)P(0, h) = -\cos(h)$ and $(\partial^2/\partial\varepsilon\partial h)P(0, h) = (-1/2)(\sin(h) + h\cos(h))$, we see that the Hessian of P at $(0, nh)$ takes the form

$$\begin{bmatrix} \frac{\partial^2}{\partial\varepsilon^2}P(0, n\pi) & \frac{n\pi}{2} \\ \frac{n\pi}{2} & 1 \end{bmatrix}.$$

(The top left entry changes with the integrator, the other three do not.) For $(0, n\pi)$ to be a minimum, the Hessian has to be positive semidefinite; since the bottom right entry is > 0, positive semidefiniteness is equivalent to nonnegative determinant, i.e. to

$$\frac{\partial^2}{\partial\varepsilon^2}P(0, n\pi) \geq \frac{n^2\pi^2}{4}.$$

However Proposition 5.1 ensures that the opposite inequality holds and we have proved the n odd case of the following result (the n even case is proved in a parallel way, changing minimum to maximum, etc.).

PROPOSITION 4.1. *Assume that an integrator of the form* (2.1) *is stable for values of* (ε, h) *in a neighbourhood of* $(0, n\pi)$*,* $n = 1, 2, \ldots$ *Then necessarily:*

$$\frac{\partial^2}{\partial\varepsilon^2}P(0, n\pi) = (-1)^{n+1}\frac{n^2\pi^2}{4}.$$

This proposition is helpful to identify suitable values of the parameters r_i and k_j in (2.1), as will be clear in our study of the stability of the families of three-stage integrators.

4.2. Palindromic methods with $m = 3$ stages. Integrators with three or fewer stages are important because, arguably, integrators with four or more stages are too complicated to be used in most applications. For the case of the kinetic/potential split systems (1.5)–(1.6), there are 3-stage integrators that clearly improve on Verlet in HMC and molecular dynamics [17, 27, 25, 2, 19, 1]. As we shall prove presently, for the (1.2)–(1.3) splitting studied in this paper, there is little room for improving on the Strang splitting. As explained in the introduction this result is very relevant when choosing the integrator for HMC algorithms to sample from target distributions resulting from perturbing a Gaussian.

For the sake of brevity we only present our findings for the K-first case in (2.1). The results for the R-first case differ in the details but yield the same conclusions. As we have noted several times, it is sufficient

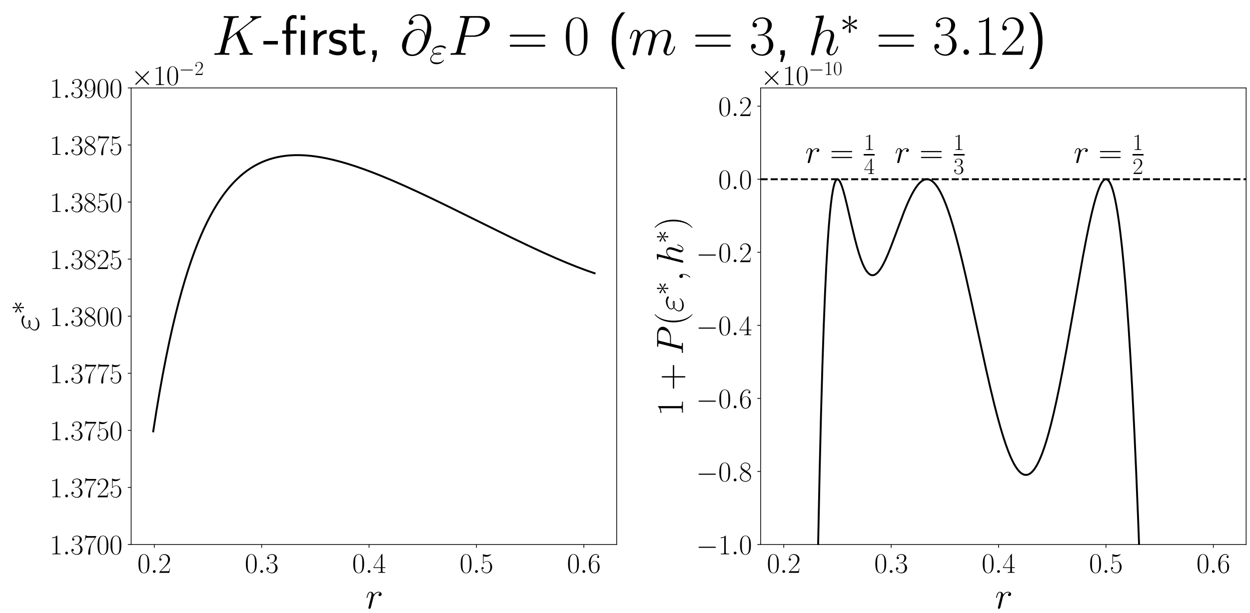

Fig. 2: Palindromic three-stage, K-first integrators. On the left, for each r, the value ε^* of the parameter ε that locally minimizes the stability polynomial $P(\varepsilon, h^*, r)$. On the right, the minimum value $P(\varepsilon^*, h^*, r)$ as a function of r: except for three exceptional cases (see text), all integrators show $P < -1$, i.e. exponential instability.

to study the palindromic case, for which, after imposing consistency, integrators take the form

$$\psi_h = \varphi^{[K]}_{kh} \circ \varphi^{[R]}_{rh} \circ \varphi^{[K]}_{(1/2-k)h} \circ \varphi^{[R]}_{(1-2r)h} \circ \varphi^{[K]}_{(1/2-k)h} \circ \varphi^{[R]}_{rh} \circ \varphi^{[K]}_{kh}. \tag{4.1}$$

There are two free parameters k and r. If we wish to have stability in a neighbourhood of $(0, \pi)$ in the (ε, h) plane, we have to impose the necessary condition in Proposition 4.1, that for (4.1) is found to read

$$4k \sin^2(\pi r) = -\cos(2\pi r).$$

However, this condition is only necessary for P to have a minimum $P = -1$ at $\varepsilon = 0$, $h = \pi$. To investigate the behaviour of P in the neighbourhood of $(0, \pi)$, we proceed as follows. We use the last display to express k in terms of r and see P as a function of (ε, h, r). We then fix a value $h^* = 3.12$ of h slightly below π and look at the behaviour of $P(\varepsilon, h^*, r)$. For each r in a suitable range,[1] we identify the value $\varepsilon^*(h^*, r) \approx 0$ of ε for which $(\partial/\partial\varepsilon)P(\varepsilon, h^*, r)$ vanishes (and therefore the function $\varepsilon \mapsto P(\varepsilon, h^*, r)$ may achieve a minimum) and plug this value into P to obtain a function $F(r) = P(\varepsilon(h^*, r), h^*, r)$ of the real variable r. This function is plotted in the right panel of Figure 2, where we see that for "most" values of r, $F(r)$ takes values below -1, indicating exponential instability of the integrator. There are however three exceptional values of r, where $F = -1$:

- $r = 1/4$. This leads to $k = 0$ so that the first and last kicks in (4.1) are the identity and may be suppressed. The integrator is then seen to be RKR2, that we know is indeed stable in the neighbourhood of $(0, \pi)$.
- $r = 1/3$. This yields KRK3, that we know is stable in the neighbourhood of $(0, \pi)$ (and also in the neighbourhood of $(0, 2\pi)$).
- $r = 1/2$. Now the central rotation in (4.1) is the identity. The integrator is KRK2, that we know is stable in the neighbourhood of $(0, \pi)$.

[1] We present results for $r \in [0.2, 0.6]$. Values of r outside this interval are not of interest as a preliminary computer search shows they have poor stability properties near $\varepsilon = -1$.

The values of ε^* where the algorithm has been found to be exponentially unstable are plotted in the left panel of Figure 2. This shows that, for $h = 3.12$, all the integrators considered (with the exceptions of RKR2, KRK2 and KRK3) are unstable for values of ε extremely close to 0. For comparison, using (2.10), one sees that for $h = 3.12$, RKR3 and KRK3 are stable for $\varepsilon \in (-1, 3.36)$ and RKR2, KRK2 are stable for $\varepsilon \in (-1, 1.30)$. Also, from Theorem 3.1, for fixed, very small $\varepsilon > 0$, RKR3 and KRK3 are stable up to $h \approx 3\pi$, while most three stage integrators have lost stability before h reaches π. The conclusion is clear: three-stage splitting integrators different from Strang have very limited stability domains.

5. A technical result.

In this section we establish the following result that was used to prove Proposition 4.1:

PROPOSITION 5.1. *The stability polynomial $P(\varepsilon, h)$ of any (consistent) splitting integrator* (2.1) *satisfies*

$$\frac{\partial^2}{\partial \varepsilon^2} P(0, n\pi) \leq \frac{n^2\pi^2}{4}, \qquad n = 1, 3, \dots,$$

and

$$\frac{\partial^2}{\partial \varepsilon^2} P(0, n\pi) \geq -\frac{n^2\pi^2}{4}, \qquad n = 2, 4, \dots$$

Proof. We recommence from (2.7) in the proof of Proposition 2.1. The coefficient of ε^2 in the right hand-side of that equality is, with $\eta_i = \sum_{n=i+1}^{m+1} r_n$, $\theta_j = \sum_{n=1}^{j} r_n$,

$$h^2 \sum_{i=2}^{m} \sum_{j=1}^{i-1} k_i k_j \exp(\eta_i hR) K \exp((1 - \eta_i - \theta_j)hR) K \exp(\theta_j hR),$$

where, by using the expressions for $\exp(tR)$ and K, the product of matrices in the summation may be computed as

$$\begin{bmatrix} -\big(\sin(h\theta_j)\cos(h(1-\eta_i-\theta_j)) - \sin(h(1-\eta_i))\big)\sin(h\eta_i) & \cdots \\ \cdots & -\big(\sin(h\eta_i)\cos(h(1-\eta_i-\theta_j)) - \sin(h(1-\theta_j))\big)\sin(h\theta_j) \end{bmatrix}.$$

We next take semitraces and recall that, from Taylor's theorem, the coefficient of ε^2 in a polynomial equals twice its second derivative evaluated at $\varepsilon = 0$. In this way we find

$$\frac{\partial^2}{\partial \varepsilon^2} P(0, h) = -h^2 \sum_{i=2}^{m} \sum_{j=1}^{i-1} k_i k_j \Big[\big(\sin(h\theta_j)\cos(h(1-\eta_i-\theta_j)) - \sin(h(1-\eta_i))\big)\sin(h\eta_i)$$
$$+ \big(\sin(h\eta_i)\cos(h(1-\eta_i-\theta_j)) - \sin(h(1-\theta_j))\big)\sin(h\theta_j) \Big].$$

By transforming the products of trigonometric functions into sums, we obtain

$$\frac{\partial^2}{\partial \varepsilon^2} P(0, h) = \frac{h^2}{2} \sum_{i=2}^{m} \sum_{j=1}^{i-1} k_i k_j \left(\cos\left(2h \left(\frac{1}{2} - \eta_i - \theta_j \right) \right) - \cos(h) \right),$$

and evaluating at $h = n\pi$ we find, after some additional trigonometric manipulations,

$$\frac{\partial^2}{\partial \varepsilon^2} P(0, n\pi) = (-1)^{n+1} n^2 \pi^2 \sum_{i=2}^{m} \sum_{j=1}^{i-1} k_i k_j \sin^2\big(n\pi(\eta_i + \theta_j)\big).$$

We now note that $\eta_i + \theta_j = 1 - (\theta_i - \theta_j)$ and $\sin^2(n\pi - (\theta_i - \theta_j)) = \sin^2(n\pi(\theta_i - \theta_j))$, so that

$$\frac{\partial^2}{\partial \varepsilon^2} P(0, n\pi) = (-1)^{n+1} n^2 \pi^2 \sum_{i=2}^{m} \sum_{j=1}^{i-1} k_i k_j \sin^2\big(n\pi(\theta_i - \theta_j)\big).$$

The proof will be ready if we prove that

$$\sum_{i=2}^{m} \sum_{j=1}^{i-1} k_i k_j \sin^2\big(n\pi(\theta_i - \theta_j)\big) \leq \frac{1}{4},$$

or, writing the double sum in a more symmetric form,

$$S = \sum_{i=1}^{m} \sum_{j=1}^{m} k_i k_j \sin^2 \big(n\pi(\theta_i - \theta_j) \big) \leq \frac{1}{2}.$$

At this point, it is convenient to assume that (i) m is even and (ii) the integrator is palindromic. As noted before there is no loss of generality in assuming (ii). And m may always be taken to be even by adding dummy stages. The double sum S may be decomposed as

$$S = \sum_{i=1}^{m} \sum_{j=1}^{m} = \sum_{i=1}^{m/2} \sum_{j=1}^{m/2} + \sum_{i=1}^{m/2} \sum_{j=m/2}^{m} + \sum_{i=m/2}^{m} \sum_{j=1}^{m/2} + \sum_{i=m/2}^{m} \sum_{j=m/2}^{m},$$

which, by symmetry, implies

$$S = 2 \sum_{i=1}^{m/2} \sum_{j=1}^{m/2} k_i k_j \sin^2 \big(n\pi(\theta_i - \theta_j) \big) + 2 \sum_{i=1}^{m/2} \sum_{j=m/2}^{m} k_i k_j \sin^2 \big(n\pi(\theta_i - \theta_j) \big).$$

and, since $k_{m+1-j} = k_j$, $\theta_i - \theta_{m+1-j} = \theta_i + \theta_j - 1$, $\sin^2 \big(n\pi(\theta_i + \theta_j - 1) \big) = \sin^2 \big(n\pi(\theta_i + \theta_j) \big)$,

$$S = 2 \sum_{i=1}^{m/2} \sum_{j=1}^{m/2} k_i k_j \sin^2 \big(n\pi(\theta_i - \theta_j) \big) + 2 \sum_{i=1}^{m/2} \sum_{j=1}^{m/2} k_i k_j \sin^2 \big(n\pi(\theta_i + \theta_j) \big).$$

We finally invoke the trigonometric identity $\sin^2(A+B) + \sin^2(A-B) = 1 - \cos(2A)\cos(2B)$ and write

$$\begin{aligned} S &= 2 \sum_{i=1}^{m/2} \sum_{j=1}^{m/2} k_i k_j \Big(1 - \cos(2n\theta_i) \cos(2n\theta_j) \Big) \\ &= 2 \left(\sum_{i=1}^{m/2} k_i \right)^2 - 2 \left(\sum_{i=1}^{m/2} k_i \cos(2n\pi\theta_i) \right)^2 \\ &\leq 2 \left(\sum_{i=1}^{m/2} k_i \right)^2 = \frac{1}{2}, \end{aligned}$$

and the proof is complete. ∎

Acknowledgments. We would like to thank Sergio Blanes for alerting us to the open problem of the Strang splitting's optimality for the alternative R, K integrators.

REFERENCES

[1] A. Ahmed and S. Adrian, *A Hybrid Monte Carlo Sampling Filter for Non-Gaussian Data Assimilation*, AIMS Geosciences, 3 (2015), pp. 41–78.

[2] M. Aleardi and A. Salusti, *Hamiltonian Monte Carlo Algorithms for Target and Interval-oriented Amplitude versus Angle Inversions*, Geophysics, 85 (2020), pp. R177–R194.

[3] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, vol. 60, Springer-Verlag, New York, 2nd ed., 1989.

[4] A. Beskos, K. Kalogeropoulos, and E. Pazos, *Advanced MCMC Methods for Sampling on Diffusion Pathspace*, Stochastic Processes and their Applications, 123 (2013), pp. 1415–1453.

[5] A. Beskos, F. J. Pinski, J. M. Sanz-Serna, and A. M. Stuart, *Hybrid Monte Carlo on Hilbert Spaces*, Stochastic Processes and their Applications, 121 (2011), pp. 2201–2230.

[6] S. Blanes and F. Casas, *A Concise Introduction to Geometric Numerical Integration*, CRC Press, 2017.

[7] S. Blanes, F. Casas, and A. Murua, *On the Linear Stability of Splitting Methods*, Foundations of Computational Mathematics, 8 (2008), pp. 357–393.

[8] S. Blanes, F. Casas, and J. M. Sanz-Serna, *Numerical Integrators for the Hybrid Monte Carlo Method*, SIAM Journal on Scientific Computing, 36 (2014), pp. A1556–A1580.

[9] N. Bou-Rabee, *Cayley Splitting for Second-order Langevin Stochastic Partial Differential Equations*, arXiv preprint arXiv:1707.05603, (2017).

[10] N. Bou-Rabee and J. M. Sanz-Serna, *Geometric Integrators and the Hamiltonian Monte Carlo Method*, Acta Numerica, 27 (2018), pp. 113–206.

[11] J. C. Butcher, *The Effective Order of Runge–Kutta Methods*, in Proceedings of the Conference on the Numerical Solution of Differential Equations, J. L. Morris, ed., vol. 109 of Lecture Notes in Mathematics, Springer, 1969, pp. 133–139.

[12] J. C. Butcher, *An Algebraic Theory of Integration Methods*, Mathematics of Computation, 26 (1972), pp. 79–106.

[13] J. C. Butcher and J. M. Sanz-Serna, *The Number of Conditions for a Runge–Kutta Method to Have Effective Order p*, Applied Numerical Mathematics, 22 (1996), pp. 103–111.

[14] C. M. Campos and J. M. Sanz-Serna, *Palindromic 3-stage Splitting Integrators, a Roadmap*, Journal of Computational Physics, 346 (2017), pp. 340–355.

[15] F. Casas, J. M. Sanz-Serna, and L. Shaw, *Split Hamiltonian Monte Carlo Revisited*, Statistics and Computing, 32 (2022), pp. 1–14.

[16] M. M. Chawla and S. R. Sharma, *Intervals of Periodicity and Absolute Stability of Explicit Nyström Methods for $y'' = f(x, y)$*, BIT Numerical Mathematics, 21 (1981), pp. 455–464.

[17] M. Fernández-Pendás, E. Akhmatskaya, and J. M. Sanz-Serna, *Adaptive Multi-stage Integrators for Optimal Energy Conservation in Molecular Simulations*, Journal of Computational Physics, 327 (2016), pp. 434–449.

[18] B. García-Archilla, J. M. Sanz-Serna, and R. D. Skeel, *Long-time-step Methods for Oscillatory Differential Equations*, SIAM Journal on Scientific Computing, 20 (1998), pp. 930–963.

[19] F. Goth, *Higher Order Auxiliary Field Quantum Monte Carlo Methods*, in Journal of Physics: Conference Series, vol. 2207, 2022, p. 012029.

[20] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration*, Springer Berlin, Heidelberg, 2nd ed., 2006.

[21] M. Hochbruck and A. Ostermann, *Exponential Integrators*, Acta Numerica, 19 (2010), pp. 209–286.

[22] R. Jeltsch and O. Nevanlinna, *Stability of Explicit Time Discretizations for Solving Initial Value Problems*, Numerische Mathematik, 37 (1981), pp. 61–91.

[23] M. G. Krein, *A Generalization of Some Investigations of A. M. Lyapunov on Linear Differential Equations with Periodic Coefficients*, Doklady Akademii Nauk SSSR, 73 (1950), pp. 445–448.

[24] B. Leimkuhler and C. Matthews, *Molecular Dynamics*, Springer International Publishing, Cham, 2015.

[25] J. Mannseth, T. S. Kleppe, and H. J. Skaug, *On the Application of Improved Symplectic Integrators in Hamiltonian Monte Carlo*, Communications in Statistics-Simulation and Computation, 47 (2018), pp. 500–509.

[26] R. M. Neal, *MCMC Using Hamiltonian Dynamics*, in Handbook of Markov Chain Monte Carlo, S. Brooks, A. Gelman, G. L. Jones, and X.-L. Meng, eds., Chapman and Hall/CRC, 2011, pp. 139–188.

[27] A. Nishimura, D. B. Dunson, and J. Lu, *Discontinuous Hamiltonian Monte Carlo for Discrete Parameters and Discontinuous Likelihoods*, Biometrika, 107 (2020), pp. 365–380.

[28] J. M. Sanz-Serna, *Markov Chain Monte Carlo and Numerical Differential Equations*, in Current Challenges in Stability Issues for Numerical Differential Equations, L. Dieci and N. Guglielmi, eds., Springer International Publishing, Cham, 2014, pp. 39–88.

[29] J. M. Sanz-Serna and M. P. Calvo, *Numerical Hamiltonian Problems*, Chapman and Hall, London, 1994.

[30] J. M. Sanz-Serna and M. N. Spijker, *Regions of Stability, Equivalence Theorems and the Courant-Friedrichs-Lewy Condition*, Numerische Mathematik, 49 (1986), pp. 319–329.

[31] B. Shahbaba, S. Lan, W. O. Johnson, and R. M. Neal, *Split Hamiltonian Monte Carlo*, Statistics and Computing, 24 (2014), pp. 339–349.

[32] G. Strang, *Accurate Partial Difference Methods I: Linear Cauchy Problems*, Archive for Rational Mechanics and Analysis, 12 (1963), p. 392–402.