Explicit geometric integration of polynomial vector fields

Robert I. McLachlan[*] and G. Reinout W. Quispel[†]

June 25, 2004

Dedicated to Syvert Nørsett on the occasion of his 60th birthday.

Abstract

We present a unified framework in which to study splitting methods for polynomial vector fields in $\mathbb{R}^n$. The vector field is to be represented as a sum of shears, each of which can be integrated exactly, and each of which is a function of $k < n$ variables. Each shear must also inherit the structure of the original vector field: we consider Hamiltonian, Poisson, and volume-preserving cases. Each case then leads to the problem of finding an optimal distribution of points on an appropriate homogeneous space, generally the Grassmannians of k-planes or (in the Hamiltonian case) isotropic k-planes in $\mathbb{R}^n$. These optimization problems have the same structure as those of constructing optimal experimental designs in statistics.

1 Introduction

The task of arranging s points on a sphere as 'evenly' as possible has an appealingly simple and fundamental air about it [19]. It is not surprising that it has been studied for a long time and has found many applications [11]. The answer, of course, depends on what we mean by 'evenly arranged'; for $s = 12$ points, for example, the 12 vertices of a regular icosahedron would satisfy most people, while for other values of s the best arrangements depend on the criterion. Maximizing the minimum interpoint distance and minimizing the potential energy of Coulomb particles placed at the points are two examples of such criteria.

We can also think of generalizing the problem while hopefully retaining its fundamental character. The dimension of the sphere can be changed, and Conway, Hardin, and Sloane [3] have considered arranging s lines (or, more generally, k-planes) through the origin in $\mathbb{R}^n$, that is, choosing s points in the Grassmannian $G(n, k)$ of k-planes. Both the sphere $\mathbb{S}^n$ and Grassmannian $G(n, k)$ are homogeneous spaces, and their homogeneity seems to be essential to the problem's appeal—arranging

---

[*]IFS, Massey University, Palmerston North, New Zealand; R.McLachlan@massey.ac.nz
[†]Department of Mathematics, La Trobe University, Bundoora, Victoria, Australia; R.Quispel@latrobe.edu.au

points on a square would not seem so important. The orthogonal group $O(n)$ acts on S^n and on $G(n,k)$, and the objective function which measures how evenly arranged the s points are is invariant under $O(n)$: two configurations related by a rotation or reflection are considered equally good. (Perhaps the fact that these homogeneous spaces carry an invariant measure, so that we have a sense of how evenly distributed the points are, is also relevant.)

In this paper we show how certain problems involving expressing polynomials as the sum of polynomials in fewer variables can be phrased as problems of finding 'good' arrangements of points on homogeneous spaces. The homogeneous spaces which arise are the Grassmannians $G(n,k)$ (especially $k = 1$ and $k = n-1$, when they are simply the projective space $\mathbb{P}^{n-1}$, so that we are arranging pairs of antipodal points on the sphere $\mathbb{S}^{n-1}$), isotropic Grassmannians $IG(n,k)$ (especially $k = n/2$, the Lagrangian Grassmannian), and tori $\mathbb{T}^n$. These problems arise when constructing explicit, geometric (structure-preserving) numerical integrators for polynomial vector fields; the isotropic and Lagrangian Grassmannians arise in the case of Hamiltonian vector fields.

To illustrate, suppose we wish to express homogeneous polynomials of degree d in n variables as the sum of s polynomials in 1 variable, i.e.

$$P(x_1, \ldots, x_n) = \sum_{i=1}^{s} c_i (a_i \cdot x)^d.$$

We let $c = (c_1, \ldots, c_s)^{\mathrm{T}} \in \mathbb{R}^s$ and $A = (a_1, \ldots, a_s) \in \mathbb{R}^{n \times s}$, and call A the 'basis'. Three approaches suggest themselves.

1. For a given P we can try and determine a suitable c, A, and minimum s for this P. This problem has a long history. The problem of determining the minimum s for a generic P is called the *algebraic Waring problem*; it was open for 120 years and has only recently been solved [1, 12]. Note that this does not determine the value of s for a specific P. Even if s is known, finding A for a given P involves solving a large set of nonlinear polynomial equations. Furthermore, as a generic P approaches the 'bad' set of non-generic polynomials, the coefficients A will blow up, which as we shall see is very bad for our application.

2. For a given P one can take a larger value of s, one sufficient to decompose all polynomials, and then adjust the basis A to obtain a "good" decomposition of P, as done by Blanes [2]. This avoids both problems: the equations to be solved are now linear, and there are no "bad" P's to avoid. However, one has to solve an optimization problem for each P, which is expensive, and much of the intrinsic symmetry of the problem is lost; there is a complicated relationship between P and the choice of A.

3. In the third approach, adopted in this paper, one takes the value of s which is sufficient for all P, and looks for a basis A which will be used for *all* polynomials P of the given degree d and dimension n. A will be adjusted so that it provides a good decomposition for all P.

For example, the 3 quadratics x^2, $(x+y)^2$, and y^2 are all pure powers (i.e., functions of $k = 1$ variable) and span the quadratics in two variables, while the 4 pure cubics x^3, $(x+y)^3$, $(x-y)^3$, and y^3 span all cubics in two variables. We shall see that these 4 cubics are the 'best possible' in a certain sense, while the 3 quadratics are not.

(In adopting the third approach, one should check that, compared to the second approach, not too much accuracy is lost on average. In the worst case, the third approach is definitely inferior. Suppose $P = x_1^d$. Then the error using a specially chosen basis, namely $s = 1$ and basis x_1, is zero.)

Suppose that P is given in a basis of monomials as

$$P(x) = \sum_{j_1 + \ldots + j_n = d} p_{j_1, \ldots, j_n} \prod_{i=1}^{n} \frac{x_i^{j_i}}{\sqrt{j_i!}}$$
$$:= \sum_{|\mathbf{j}| = d} p_{\mathbf{j}} \frac{x^{\mathbf{j}}}{\sqrt{\mathbf{j}!}}$$

Recall that the number of monomials of degree d in n variables, i.e. the dimension of the space of such homogeneous polynomials, is $N(n, d) := \binom{n-1+d}{d}$. It is $\mathcal{O}(n^d)$ as $n \to \infty$ for fixed d, and $\mathcal{O}(d^{n-1})$ as $d \to \infty$ for fixed n. We collect all $N(n, d)$ coefficients $p_{j_1, \ldots j_n}$ into a vector p.

The basis monomials $x^{\mathbf{j}} / \sqrt{\mathbf{j}!}$ are used because this makes the norm $\|P\| := \|p\|_2$ rotationally invariant—a crucial part of the problem of distributing points on spheres, for example. (See also [6].)

Proposition 1 Let $P(x) = \sum p_{\mathbf{j}} \frac{x^{\mathbf{j}}}{\sqrt{\mathbf{j}!}} = Q(y) = \sum q_{\mathbf{j}} \frac{y^{\mathbf{j}}}{\sqrt{\mathbf{j}!}}$ where $x = Ry$ and $R \in \mathrm{O}(n)$. Then $\|p\|_2 = \|q\|_2$.

Proof The result is clearly true for $R = -I$, so it remains to check it for $R \in \mathrm{SO}(n)$. Let $R = \exp(tA)$ where $A \in \mathfrak{so}(n)$. Then

$$\dot{q}|_{t=0} = \sum_i \frac{\partial Q}{\partial x_i} \dot{x}_i$$
$$= \sum_{ij} \frac{\partial Q}{\partial x_i} A_{ij} x_j$$
$$= \sum_{\mathbf{l}ij} q_{\mathbf{l}} l_i A_{ij} x_j x^{\mathbf{l}} \frac{1}{x_i \sqrt{\mathbf{l}!}}$$
$$= \sum_{\mathbf{m}ij} q_{\mathbf{m}+e_i-e_j} (m_i + 1) A_{ij} \frac{1}{\sqrt{(\mathbf{m} + e_i - e_j)!}} x^{\mathbf{m}}$$

where e_i is the ith elementary unit vector, so we have

$$\frac{d}{dt}\bigg|_{t=0} \frac{1}{2}\|q\|_2^2 = q \cdot \dot{q}|_{t=0}$$

$$= \sum_{\mathbf{m}ij} q_{\mathbf{m}} q_{\mathbf{m}+e_i-e_j}(m_i+1)A_{ij}\frac{\sqrt{\mathbf{m}!}}{\sqrt{(\mathbf{m}+e_i-e_j)!}}$$

$$= \sum_{\mathbf{m}ij} q_{\mathbf{m}} q_{\mathbf{m}+e_i-e_j}A_{ij}\sqrt{(m_i+1)m_j}$$

$$= \sum_{\mathbf{l}ij} q_{\mathbf{l}-e_i+e_j} q_{\mathbf{l}}A_{ij}\sqrt{l_i(l_j+1)} \quad (\text{setting } \mathbf{l}=\mathbf{m}+e_i-e_j)$$

$$= -\sum_{\mathbf{m}ij} q_{\mathbf{m}+e_i-e_j} q_{\mathbf{m}}A_{ij}\sqrt{m_j(m_i+1)} \quad (\text{setting } \mathbf{l}=\mathbf{m}, i \leftrightarrow j)$$

$$= 0,$$

using $A^{\mathrm{T}} = -A$. Since $\|q\|_2 = \|p\|_2$ at $t=0$, the result is proved. $\bullet$

We now have to match coefficients of x in

$$\sum_{\mathbf{j}} p_{\mathbf{j}}\frac{1}{\sqrt{\mathbf{j}!}}x^{\mathbf{j}} = \sum_{i=1}^{s} c_i \frac{1}{\sqrt{d!}}(a_i \cdot x)^d$$

$$= \sum_{i,\mathbf{j}} c_i \frac{\sqrt{d!}}{\mathbf{j}!}a_i^{\mathbf{j}}x^{\mathbf{j}}$$

which gives the linear system

$$Bc = p \quad \text{where } B_{\mathbf{j},i} = \frac{\sqrt{d!}}{\sqrt{\mathbf{j}!}}a_i^{\mathbf{j}}. \tag{1}$$

We shall see below that we want to have $\|c\|$ as small as possible. We shall consider only the 2-norm of c (see, e.g., the discussion on choice of norm in [11]).

Let $B = U\Sigma V^{\mathrm{T}}$ be the singular value decomposition of B. Then

$$\Sigma\tilde{c} = \tilde{p}$$

where $c = V\tilde{c}$ and $p = U\tilde{p}$. Assuming that B has full rank, the singular values σ_i of B are nonzero and the solution to (1) of minimum 2-norm is therefore $\tilde{c}_i = \tilde{p}_i/\sigma_i$.

That is, each choice of s vectors a_i determines an expansion of P as a sum of polynomials in 1 variable. By scaling c_i and a_i, only the direction of a_i is important, i.e. we have to choose s lines through the origin in $\mathbb{R}^n$, or s points in the real projective space $\mathbb{P}^{n-1}$.

To find bases which are good on average for all polynomials P, we let p range over the unit sphere $\|p\|_2 = 1$. Then $\|c\|_2$ ranges over $[1/\max\sigma_i, 1/\min\sigma_i]$. The average value of $\|c\|_2^2$ is $\|1/\sigma\|_2^2/N$,

where $N := N(n, d)$ is the number of singular values (= number of rows of B). For, this average is

$$\frac{\int_{\|p\|=1} \|c\|^2 \mathrm{d}\mu}{\int_{\|p\|=1} \mathrm{d}\mu} = \frac{\int_{\|p\|=1} \sum_{i=1}^{N} p_i^2/\sigma_i^2 \, \mathrm{d}\mu}{\int_{\|p\|=1} \mathrm{d}\mu}$$

$$= \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \frac{\int_{\|p\|=1} p_i^2 \mathrm{d}\mu}{\int_{\|p\|=1} \mathrm{d}\mu}$$

$$= \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \frac{\int_{\|p\|=1} \frac{1}{N} \sum_{j=1}^{N} p_j^2 \mathrm{d}\mu}{\int_{\|p\|=1} \mathrm{d}\mu}$$

$$= \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \frac{\int_{\|p\|=1} \mathrm{d}\mu}{\int_{\|p\|=1} \mathrm{d}\mu}$$

$$= \frac{1}{N} \|1/\sigma\|_2^2$$

where $\mathrm{d}\mu$ is an invariant measure on the sphere $\|p\| = 1$. We define the *condition number* of the basis A by

$$\kappa(A) := \frac{1}{\sqrt{N}} \|1/\sigma\|_2.$$

Note that $\kappa(A)^2 = \frac{1}{N} \operatorname{tr}((BB^{\mathrm{T}})^{-1})$.

Example 1 *The simplest conceivable case is that of quadratics in 2 variables. We need to choose $s = 3$ lines through the origin in $\mathbb{R}^2$. Let the angles of these lines be θ_1, θ_2, and θ_3, and choose the basis of monomials as $\{\frac{1}{\sqrt{2}}x_1^2, x_1 x_2, \frac{1}{\sqrt{2}}x_2^2\}$. Then we find*

$$B = \begin{pmatrix} \sqrt{2}\cos^2\theta_1 & \sqrt{2}\cos^2\theta_2 & \sqrt{2}\cos^2\theta_3 \\ 2\cos\theta_1\sin\theta_1 & 2\cos\theta_2\sin\theta_2 & 2\cos\theta_3\sin\theta_3 \\ \sqrt{2}\sin^2\theta_1 & \sqrt{2}\sin^2\theta_2 & \sqrt{2}\sin^2\theta_3 \end{pmatrix}.$$

If we take $\theta = (0, \pi/4, \pi/2)$ (corresponding to the basis $\{x_1^2, (x_1+x_2)^2, x_2^2\}$, as we used for illustration above), then the inverse singular values of B are $(1.31, 0.71, 0.54)$, and $\kappa(A) \approx 0.91$.

On the other hand, if we take $\theta = (0, \pi/3, 2\pi/3)$ (corresponding to the basis $\{x_1^2, (\frac{1}{2}x_1+\frac{\sqrt{3}}{2}x_2)^2, (-\frac{1}{2}x_1+\frac{\sqrt{3}}{2}x_2)^2\}$), the inverse singular values of B are $(0.82, 0.82, 0.58)$ and $\kappa(A) \approx 0.75$. For almost all polynomials, all coefficients c_i will be smaller for this basis than for the previous one. One can show that this configuration (or any rotation of it) achieves the minimum of κ over all θ.

There are also nearly singular configurations. At $\theta = (0, 0.1, 0.2)$ the reciprocal singular values of B are $(86.32, 3.56, 0.41)$. Only polynomials which happened to lie in the correct singular subspace would be well represented. Such θs are the 'bad sets' of basis vectors which we need to stay away from. Achieving $1/\sigma_i < 1$ for all i, as we can in this case, is nice but is not generally possible.

For a specific P, there are indeed slightly better bases. For example, for $P = x_1 x_2$, the above 'optimal' basis has $c = \frac{1}{\sqrt{3}}(0, 1, 1)$, while the basis $\theta = (0, \pi/4, 3\pi/4)$ has $c = \frac{1}{2}(0, 1, 1)$, which is slightly smaller.

This leads to the problem of determining the basis A which minimizes the condition number $\kappa(A)$ for each n, d, and s.

2   Application to numerical integrators

The application to numerical integrators arises because some vector fields which are functions of fewer than n variables are easy to integrate exactly; and exact flows of vector fields inherit the geometric properties of the vector fields, such as symplecticity. So such vector fields are useful ways of constructing (e.g.) symplectic maps. See [16] for a review of splitting methods.

They are also useful for constructing symplectic maps which approximate a given nearly symplectic map; it is in this context that the splitting theory for polynomials has been largely developed [6, 7, 8, 13, 18].

For Poisson (e.g. Hamiltonian) systems, the following gives a large set of easily integrated systems which can be used for splitting. On a Poisson manifold $(M, \{,\})$, the Hamiltonian vector field associated to the Hamiltonian H is $X_H := \{\cdot, H\}$. Equivalently, we can specify a Poisson 2-vector J, so that $\{F, G\}(x) = J(x)(dF, dG)$ and $\dot{x} = J(x)(\cdot, dH)$.

Proposition 2 *[16] Let $F_1, \ldots, F_k$ be k functions such that X_{F_i} is integrable for all j and $\{F_i, F_j\} = 0$ for all i and j. Then, X_H for $H = H(F_1, \ldots, F_k)$ is integrable. Furthermore, if $\frac{\partial H}{\partial F_i}$ and $\exp(X_{F_i})$ can be evaluated in terms of elementary functions, then so can $\exp(X_H)$, which is equal to the composition of the flows of the X_{F_i} for times $\frac{\partial H}{\partial F_i}$, i.e. $\exp(X_{F_i}) = \prod_{i=1}^{k} \exp\left(\frac{\partial H}{\partial F_i} X_{F_i}\right)$.*

We now introduce some classes of polynomial vector fields and ways of splitting them that lead to choosing sets of points on different manifolds.

1. Let J be an $n \times n$ antisymmetric matrix and consider the Poisson system

$$\dot{x} = J\nabla P(x). \tag{2}$$

   In the case $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$, this is a canonical Hamiltonian system. Suppose for some s we can find $A \in \mathbb{R}^{n \times s}$ and functions $P_1, \ldots P_s$ such that

$$P(x_1, \ldots, x_n) = \sum_{j=1}^{s} P_j\left(a_j^{\mathrm{T}} x\right).$$

   Then the Poisson system $\dot{x} = J\nabla P_j$ has the form

$$\dot{x} = P_j'(a_j^{\mathrm{T}} x) J a_j \tag{3}$$

   where a_j is the jth column of A. Then $a_j^{\mathrm{T}} \dot{x} = P_j'(a_j^{\mathrm{T}} x)(a_j^{\mathrm{T}} J a_j) = 0$, so $a_j^{\mathrm{T}} x = $ const. and the solution of (3) is $x(t) = x(0) + t P_j'(a_j^{\mathrm{T}} x(0)) J a_j$. That is, the Hamiltonian vector field (3)

is a 'shear', a vector field on $\mathbb{R}^n$ which in suitable coordinates (y, z) takes the form $\dot{y} = 0$, $\dot{z} = f(y)$. Shears are an important subclass of the class of the 'nilpotent of degree 2' vector fields X, which have linear flows, i.e. $\exp(tX)(x) = x + tX(x)$, and are integrated exactly by Euler's method [9, 16].

The flow of (2) is then typically approximated by a symplectic map such as the leapfrog method

$$\prod_{i=1}^{s} \exp(\frac{1}{2}\Delta t J\nabla P_i) \prod_{i=s}^{1} \exp(\frac{1}{2}\Delta t J\nabla P_i)(x(0)) = \exp(\Delta t J\nabla P)(x(0)) + \mathcal{O}((\Delta t)^3). \quad (4)$$

When P is a homogeneous polynomial of degree d, P_j must also be a homogeneous polynomial of degree d. This is the case considered in the previous section. It requires finding 'good' sets of s points in $\mathrm{G}(n, 1) = \mathbb{P}^{n-1}$.

2. If we split P into a sum of functions P_i of more than 1 variable, the resulting Poisson systems $\dot{x} = J\nabla P_i(x)$ may not be integrable. Consider the Hamiltonian $Q(y_1, \ldots, y_k)$ where $y_i = a_i^\mathrm{T} x$. Then

$$\dot{y}_i = a_i^\mathrm{T}\dot{x} = a_i^\mathrm{T}J\nabla Q = \sum_{j=1}^{k} a_i^\mathrm{T}Ja_j \frac{\partial Q}{\partial y_j}.$$

This is a Poisson system with (constant) Poisson structure matrix $\widetilde{J}_{ij} := a_i^\mathrm{T}Ja_j$. The rank of $\widetilde{J}$ is even; if it is equal to 2, then the ODEs for y_i are integrable, the integrals being the Hamiltonian Q and the Casimirs of $\widetilde{J}$. Note that if $k = 2$ or $k = 3$, then the rank of $\widetilde{J}$ is generically equal to 2. Thus, good sets of s points in $\mathrm{G}(n, 2)$ or $\mathrm{G}(n, 3)$ generate good splittings of Hamiltonian vector fields into integrable Hamiltonian vector fields.

3. In item 2, the integrable vector fields are not very easy to integrate when the rank of $\widetilde{J}$ is 2. Therefore we now consider the case that the rank of $\widetilde{J}$ is 0. Since $\widetilde{J}$ is antisymmetric, $\widetilde{J} = 0$. Then $\dot{y}_i = 0$, $y_i = \mathrm{const.}$, $\dot{x} = JA^\mathrm{T}\nabla_y Q(y) = \mathrm{const.}$, and $\ddot{x} = 0$. That is, the Hamiltonian vector field of $P_i(x) = Q(y)$ is integrated exactly by Euler's method. We will see that the requirement $\widetilde{J} = 0$ is equivalent to the vectors a_i determining an isotropic subspace of the symplectic vector space $\mathbb{R}^{2n}$. The set of k-dimensional isotropic subspaces forms the isotropic Grassmannian $\mathrm{IG}(n, k)$. Thus, good sets of s points in $\mathrm{IG}(n, k)$ generate good splittings of Hamiltonian vector fields into Hamiltonian vector fields which are functions of k variables and are integrated exactly by Euler's method. The case $k = n/2$ is studied in [18, 7, 13].

4. Splitting is crucial in finding integrators for non-canonical Poisson systems, because there are no general Poisson integrators. Proposition 2 raises the question of finding commuting sets of integrable Hamiltonians on M, a tall order—one can regard Darboux's theorem as equivalent to constructing these functions locally. Note, however, that if we can find *any* integrable Hamiltonians F_i, not in involution, then we can at least construct explicit Poisson integrators for $H = \sum_i H_i(F_i)$.

In item 1, the Poisson structure is constant, and the linear functions $F_i = a_i^\mathrm{T} x$ are integrable; in fact, their Hamiltonian vector fields are shears. This gives enough F_i to split any polynomial Hamiltonian.

The linear functions are also integrable in the important Lie-Poisson case, in which $M = \mathfrak{g}^*$ is the dual of a Lie algebra $\mathfrak{g}$, and $\{F, G\} = \langle x, [\mathrm{d}F, \mathrm{d}G]\rangle$. (Crucially, $J(x)$ is linear in x.) Because the vector field $X_{\langle x,c\rangle}$ is linear in x for all $c \in \mathfrak{g}$ and $[X_{\langle x,c\rangle}, X_{\langle x,d\rangle}] = X_{\langle x,[c,d]\rangle}$,

7

Lie-Poisson manifolds carry a finite-dimensional Lie algebra of linear (and hence integrable) vector fields. This gives explicit Lie-Poisson integrators for $H = \sum H_i(\langle x, c_i \rangle)$, in particular, for all polynomial Lie-Poisson systems.

Moreover, let $\mathfrak{g}_1$ be an abelian subalgebra of $\mathfrak{g}$, corresponding to the space of commuting Hamiltonians $\langle x, \mathfrak{g}_1 \rangle$. From Proposition 2, to integrate $H(\langle x, \mathfrak{g}_1 \rangle)$ we only need to be able to integrate fixed linear vector fields for arbitrary times. (This is an improvement over [15], which involved integrating linear vector fields containing parameters.) The set of k-dimensional abelian subalgebras of $\mathfrak{g}$ forms a manifold, which by analogy we call the isotropic Grassmannian of $\mathfrak{g}$, $\mathrm{IG}(\mathfrak{g}, k)$. Good sets of s points in $\mathrm{IG}(\mathfrak{g}, k)$ generate good splittings of Lie-Poisson vector fields into linear vector fields. This is an interesting geometric optimization problem which (mainly because it depends on $\mathfrak{g}$) we have not explored further.

5. We also consider volume-preserving flows (see section 6). Let the vector field $\dot{x} = f(x)$ be divergence-free, $\nabla \cdot f(x) \equiv 0$, so that its flow preserves Euclidean volume. How can we split it into a sum of divergence-free, integrable vector fields? One way is to choose a vector $a_1 \in \mathbb{R}^n$, extend a_1 to an orthonormal basis $a_1, \dots, a_n$ of $\mathbb{R}^n$; then the vector field $\dot{x} = a_1 g(a_2^{\mathrm{T}} x, \dots, a_n^{\mathrm{T}} x)$ is volume preserving and a shear for any scalar function $g \colon \mathbb{R}^{n-1} \to \mathbb{R}$. When f is a polynomial of degree d, we need to choose enough vectors a_1 so that these vector fields span all divergence-free polynomial vector fields. As in case 1, we are looking for good sets of s points in $\mathrm{G}(n, 1)$, although with a different criterion of goodness.

2.1   Importance of well-conditioned bases

In this section we consider splitting a Hamiltonian into a sum of functions of k variables. We conducted a series of trials in which we chose random polynomials (uniformly on the sphere $\|p\| = 1$) and random bases A (defined by s independent points on $\mathrm{IG}(n, k)$). For each such configuration we computed the condition number $\kappa(A)$ and an estimate of the local truncation error given by the mean square energy error when Hamilton's equations $\dot{x} = J \nabla P(x)$ (J canonical) are solved numerically by splitting combined with leapfrog. The mean energy error is estimated by averaging over a large number of initial conditions on the sphere $\|x\| = 1$. The numerical evidence for $n = 4$, $d = 3$, $k = 2$, and $s = 5$ (see Figure 1) suggests that the error strongly depends on $\|c\|$, being approximately $\|c\|^z$ with $2 < z < 3$; the exponent seems to depend on P. The principal local error is a linear combination of all triple (Poisson) brackets of the P_i, giving an $\mathcal{O}(\|c\|^3)$ contribution. A possible explanation for the behaviour actually observed is related to the small value of s ($s = 5$) in our experiment. Consider sets of 2 or 3 nearly linearly dependent subspaces. This leads to large coefficients c_i for those 2 or 3 pieces in the splitting, but since the subspaces nearly Poisson-commute, a large contribution can only come from a triple Poisson bracket with one of these subspaces omitted, which drops one power of c.

So to make the splitting technique viable, we need to find configurations A in which $\kappa(A)$ is as small as possible. An absolute minimum is not required, but the nearly singular configurations should definitely be avoided. The singular configurations are codimension 1 and the smallest singular value of B measures the distance from singularity; therefore if we simply choose random configurations with respect to a uniform probability measure, the probability distribution of κ has a long tail $p(\kappa) = \mathcal{O}(\kappa^{-2})$ (because $\int_\kappa^\infty p(\kappa) d\kappa = \mathcal{O}(\kappa^{-1})$); the expected value $E[\kappa] = \int_0^\infty \kappa p(\kappa) \, d\kappa = \infty$. On the other hand, finding the global minimum of κ is nice because it essentially closes the problem,
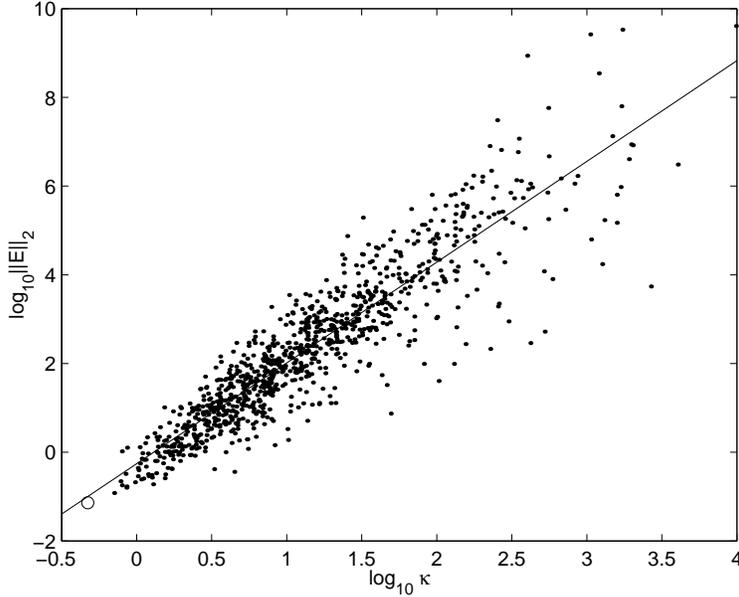
Figure 1: Dependence of integration error on the condition number of the basis. The local truncation error of the energy is shown averaged over 20 cubics in 4 variables and 100 initial conditions, for each of 1000 random bases of 5 points in $\mathrm{IG}(4,2) = \mathrm{LG}(4)$, vs. the condition number κ of the basis. The line has slope 2.27, and the circle indicates the basis of minimum κ that we have found (see Section 5).

at least for that value of n and d. For example, for $n = 3$ and $d = 2$ we need to choose 6 lines in $\mathbb{R}^3$; the optimum configuration is the lines through the 6 pairs of antipodal points of the icosahedron.

2.2   Connection with optimal design

Consider the following problem in the optimal design of experiments [10, 11]: we are to choose s points a_i on the unit sphere in $\mathbb{R}^n$ (the "design region") and fit a homogeneous polynomial $P(x) = \sum c_{\mathbf{j}} a^{\mathbf{j}}$ of degree d (the "model") to data y_i measured at a_i with independent normally distributed errors of mean 0 and variance σ^2. The least squares equations are

$$X^{\mathrm{T}} c = y \tag{5}$$

where $X_{\mathbf{j}i} = a_i^{\mathbf{j}}$. The optimal design problem is to choose the points a_i such that the fit is as reliable as possible, in some sense. The covariance matrix of the parameters is $Z := (XX^{\mathrm{T}})^{-1}$. Two popular design criteria are A-optimality, in which $\operatorname{tr} Z$ is minimized, and I-optimality, in which $\operatorname{tr} MZ$ is minimized, where M is the matrix of moments of the basis functions over a chosen region R (the "modeling region") on which the model is to be used, i.e. $M_{\mathbf{jk}} = \int_R a^{\mathbf{j}} a^{\mathbf{k}} \, da$. ($I$-optimal designs minimize the average variance of the error of the model over the modeling region.) More generally, for any matrix W of weights let us define the I_W-value of a design $A = (a_1, \ldots, a_s)$ by $I_W(A) = \operatorname{tr} WZ$.

Therefore the polynomial splitting problem is precisely an instance of the optimal design problem,

9

with a certain diagonal matrix of weights. Comparing (1) and (5), we have $B = CX$ with $C_{\mathbf{jk}} = (d!/\sqrt{\mathbf{j}!})\delta_{\mathbf{jk}}$, so $\operatorname{tr}(BB^{\mathrm{T}})^{-1} = \operatorname{tr} C^{-2}(XX^{\mathrm{T}})^{-1} = \operatorname{tr} C^{-2}W$. Therefore $\kappa(A) = \sqrt{I_{C^{-2}}(A)}/N$, and a set of points has minimum κ iff it is $I_{C^{-2}}$-optimal. The theory of optimal designs is very well studied and it is not surprising that many of the techniques of that field are useful here too. For example, a design with the minimum possible number of points s, equal to the number of parameters in the model, is called a *saturated design*. Although these have been studied, they are not very useful in practical experimental design because they are not very robust and they tend to have somewhat poor I-values. But for us, the increased cost that comes from increasing s is unlikely to compensate for having smaller condition numbers, which only decrease as $\mathcal{O}(1/\sqrt{s})$.

3 Hamiltonian case: Functions of 1 variable

From the discussion in Section 1, a basis requires us to choose at least $s = N(n,d)$ points in $\mathbb{P}^{n-1}$. On the other hand, Theorem 16 of [16] constructs a basis with exactly $s = N(n,d)$ points. This leaves the optimization problem of determining the optimal basis for each n and d.

It looks like a difficult optimization problem because the singular configurations (with $\kappa = \infty$) are codimension 1; they divide the search space into cells, and a local search cannot move from cell to cell. (Experiments with low-dimensional cases show these cells vividly; each has a single local minimum, with several nearly-optimum cells.) On the other hand, the success reported in [10] in constructing optimal designs with the program GOSSET suggests that the same approach will work here.

For many choices of n and d the optimal configurations turn out to have some symmetry. The objective function is invariant under permutations and simultaneous rotations of the s points. A study of the symmetry groups and orbit structure of some of the solutions will allow one to restrict searches to sets of points with the chosen symmetry. However, our experiments so far have suggested that extreme care is required, for many symmetries that seem reasonable turn out to lead to singular configurations, i.e., configurations that do not span all monomials. The example of $n = 3$, $d = 2$, for which the 6 optimal points are the lines joining opposite vertices of the icosahedron, suggests looking for configurations on which a discrete subgroup of $\mathrm{SO}(n)$, such as a group of reflections, acts transitively.

It would be nice if instead of carrying out an optimization for each n and d, we could find a family of solutions which was close to optimal for *all* n and d. In this section we present one approach which, while still requiring some numerical optimization, gives a solution for all n for a given d and searches over a drastically smaller space. It has the further feature that the basis for functions of n variables is the same as the basis for those functions of $n+1$ variables that are independent of x_{n+1}. (That is, the decompositions of P and Q are the same when $P(x_1, \ldots, x_{n+1}) \equiv Q(x_1, \ldots, x_n)$.) It also has full permutation invariance. This approach, together with examples of such bases, is due to Scully [20].

In this section we fix the degree d and vary n. We arrange the monomials by number of variables: first all those which are functions of 1 variable (x_i^d), then those which are functions of 2 variables ($x_i^{d-1}x_j$), and so on. We then build the basis hierarchically by first choosing the best permutation-

invariant basis for the functions of 1 variable, then choosing the best permutation-invariant basis for the functions of 2 variables, given the choice already made for the functions of 1 variable, and so on.

We illustrate the construction by working through the cases $d = 2, 3$, and 4; for each d we have to consider $n = 2, \ldots, d$.

$d = 1$ $n = 1$ For 1 variable there is no freedom and we must take the basis x_1.

         $n > 1$ When extended to n variables with permutation symmetry, the basis must include $x_1, \ldots, x_n$.

For 2 variables there are $d - 1$ monomials $(x_1^{d-1} x_2, \ldots, x_1 x_2^{d-1})$ to consider. We need to find a good basis for the $d + 1$ functions of ≤ 2 variables which includes x_1 and x_2 (inherited from the functions of 1 variable). Let the basis be $\{x_1 \cos \theta_i + x_2 \sin \theta_i : 1 \leq i \leq d + 1\}$, where $\theta_1 = 0$ and $\theta_2 = \pi/2$ are fixed.

$d = 2$ $n = 2$ Now we see how the hierarchical solution loses efficiency. The best we can do is choose $\theta_3 = \pi/4$, whereas the best solution *without* imposing the hierarchical structure is to take $\theta = (0, \pi/3, 2\pi/3)$ (Example 1).

         $n > 2$ However, in return we get a natural basis for quadratics in any number of variables, namely $\{x_i^2, (\frac{1}{\sqrt{2}}(x_i + x_j))^2\}$. For large n the corresponding points are not very equally distributed over the sphere.

$d = 3$ $n = 2$ In this case the best hierarchical and non-hierarchical solutions coincide, namely $\theta = (0, \pi/4, \pi/2, 3\pi/4)$. This basis is $\{x_1^3, x_2^3, (\frac{1}{\sqrt{2}}(x_1 \pm x_2))^3\}$.

         $n \geq 3$ To get a basis for cubics in any number of variables we now only need to consider the cubics in precisely 3 variables, of which there is only 1, namely $x_1 x_2 x_3$; to include this in the basis we need to choose a single point in $\mathbb{P}^2$, given that the 6 points $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ (from functions of 1 variable) and $(1,1,0)/\sqrt{2}$, $(1,0,1)/\sqrt{2}$, and $(0,1,1)/\sqrt{2}$ have already been chosen. It is not hard to guess that $(1,1,1)/\sqrt{3}$, the only choice with full permutation symmetry, is optimal, and this turns out to be the case.

$d = 4$ $n = 2$ We have to choose 3 angles—it is reasonable to expect that the best solution will have 1 angle in the 1st quadrant and 2 roughly equally spaced angles in the 2nd quadrant, or vice versa. Indeed, the solution with minimum κ turns out to be $\theta = \pi(0.25, 0.6524, 0.8476)$. Again, the optimal solution has full permutation invariance. When extended to functions of any number of variables, the extended basis inherits this permutation invariance. We could have instead imposed this invariance by searching amongst bases of the form $\{\pi/4, \pi/2 + \theta, \pi - \theta\}$, with a single parameter θ. Note that this set can be partitioned so that the symmetry of each set corresponds to the symmetries of the monomials, i.e. $\pi/4$ corresponds to $x_1^2 x_2^2$, and $\pi(0.6524, 0.8476)$ corresponds to $\{x_1 x_2^3, x_1^3 x_2\}$.

         $n = 3$ For quartics in $n = 3$ variables there are 3 new monomials to consider, $\{x_1^2 x_2 x_3,$ $x_1 x_2^2 x_3, \ x_1 x_2 x_3^2\}$, so we need to choose 3 points in $\mathbb{P}^2$ given that the 3 points

corresponding to x_i^4, and the 9 points corresponding to $x_i^3 x_j$ and $x_i^2 x_j^2$ have already been chosen. The optimal solution is $\{(a,b,b),(b,a,b),(b,b,a)\}$ for $a = 0.4597$, $b = -0.6280$ ($\kappa = 0.36$); that is, the optimum solution turns out to have permutation symmetry. This hierarchical solution is not that bad; the best non-hierarchical solution we have found, optimizing over 15 arbitrary points in $\mathbb{P}^2$, has $\kappa = 0.29$. Imposing permutation symmetry has reduced the dimension of the optimization problem from 30 to 1.

$n \geq 4$ Finally, to get a basis for quartics in any number of variables we add the point $(x_1 + x_2 + x_3 + x_4)/2$ and its permutations.

Regardless of whether the optimal solutions are always permutation invariant, we shall now impose permutation invariance and study its structure. We illustrate in the case $d = 5$. The numbers of monomials of degree $d = 5$ in $n \leq 1, \ldots, 5$ variables are 1, 6, 21, 56, and 126. Therefore to determine the optimum configuration in each case requires finding this number of points on $\mathbb{P}^{n-1}$ (lines through the origin in $\mathbb{R}^n$)—for example there are $4 \cdot 126 = 504$ parameters for $n = 5$. But there are only 1, 4, 6, 4, and 1 monomials of degree 5 in *precisely* 1,...,5 variables. We then divide each of these sets of monomials into their conjugacy classes under permutation, of which there are only 1, 2, 2, 1, and 1: representatives are x_i^5, $\{x_i^3 x_j^2, x_i^4 x_j\}$, $\{x_i^3 x_j x_k, x_i^2 x_j^2 x_k\}$, $x_i^2 x_j x_k x_l$, and $x_i x_j x_k x_l x_m$. Each conjugacy class is associated with a partition of d into exactly n parts. For each n we have to choose a point in $\mathbb{P}^{n-1}$ for each conjugacy class. However, the points are further constrained to inherit the symmetry of their conjugacy class; each point lies in $\mathbb{P}^{m-1}$ where m is the number of distinct powers in the conjugacy class. Therefore in the present example we have to solve optimization problems in $\mathbb{P}^0$, $\mathbb{P}^1 \times \mathbb{P}^1$, $\mathbb{P}^1 \times \mathbb{P}^1$, $\mathbb{P}^1$, and $\mathbb{P}^0$, of dimensions 0, 2, 2, 1, and 0. The total number of free parameters is reduced from 504 to 5. A better measure of complexity is perhaps the dimension of the largest optimization problem, which is 2 in this example.

We can now easily compute the sizes of the optimization problems for each d and n, namely, the sum of (the number of distinct parts of each partition of d into n parts-1). For example, for $d = 8$ and $n = 3$, we have $8 = 6 + 1 + 1 = 4 + 2 + 2 = 3 + 3 + 2 = 5 + 2 + 1 = 4 + 3 + 1$, so we have to optimize a function on $(\mathbb{P}^1)^3 \times (\mathbb{P}^2)^2$, of dimension 7. The number of partitions is a classical function, although the division of the conjugacy classes by number of distinct parts is not so common—we have simply worked these out by by listing all partitions (see Table 1). The total number of free parameters for $d = 2, \ldots, 8$ is 0, 1, 2, 5, 8, 15, and 23, and the dimension of the largest optimization problem is 0, 1, 2, 2, 3, 5, 7. It still remains to check that there in fact is a nonsingular configuration with this high degree of symmetry for each d and each $n \leq d$, and to find the optimal configurations.

4   Hamiltonian case: the Lagrangian Grassmannian

We now concentrate on the canonical case $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$ with linear functions F_i. The requirement that $J(\mathrm{d}F_i, \mathrm{d}F_j) = 0$ for all i and j is equivalent to the condition that the subspace $F := \mathrm{span}(F_1, \ldots, F_k) \subset (\mathbb{R}^n)^*$ is isotropic—that is, that the Poisson 2-vector J vanishes on it. This is equivalent to the condition that the annihilator $F^0 := \{v : f(v) = 0 \text{ for all } f \in F\} \subset \mathbb{R}^n$ is coisotropic, that is, that its symplectic orthogonal $\mathrm{orth}\, F^0$ is isotropic. Choosing a set of linear

d	n							
	1	2	3	4	5	6	7	8
2	0	0						
3	0	1	0					
4	0	1	1	0				
5	0	2	2	1	0			
6	0	2	3	2	1	0		
7	0	3	5	4	2	1	0	
8	0	3	7	6	4	2	1	0

Table 1: Size of optimization problems with hierarchical symmetry. The entry at position (n, d) gives the dimension of the associated optimization problem.

functions $F_1, \ldots F_k$ is equivalent to choosing an isotropic subspace of a symplectic vector space. The set of all isotropic subspaces forms the *isotropic Grassmannian* of $\mathbb{R}^n$ [17, 14]. It is a compact manifold $\mathrm{IG}(n, k)$ homeomorphic to $\mathrm{U}(\tilde{n})/(\mathrm{O}(k) \times \mathrm{U}(\tilde{n} - k))$, where $\tilde{n} := n/2$ and $1 \leq k \leq \tilde{n}$. It has dimension

$$\dim \mathrm{IG}(n, k) = nk - \frac{1}{2}k(3k - 1).$$

For example, $\dim \mathrm{IG}(n, 1) = n - 1$ since all 1-dimensional subspaces are isotropic and $\mathrm{IG}(n, 1) = \mathbb{P}^{n-1}$; this is the case considered in Section 3. For a given n, the isotropic Grassmannian of maximum dimension has dimension $\lfloor \frac{1}{3}\tilde{n}(2\tilde{n} + 1) \rfloor$, which is achieved by $k = \lfloor \frac{4n+1}{6} \rfloor$ for $\tilde{n} = 0 \bmod 3$, $k = \lceil \frac{4n+1}{6} \rceil$ for $\tilde{n} = 1 \bmod 3$, or by both these values of k for $\tilde{n} = 2 \bmod 3$.

For a given k we express $P : \mathbb{R}^n \to \mathbb{R}$ as a function of k variables.

Conjecture 1 *For all n, d, and k ($1 \leq k \leq n/2$), a generic choice of s^* points in $\mathrm{IG}(n, k)$ yields a basis for the polynomials of degree d, where*

$$s^* = \lceil N(n, d)/N(k, d) \rceil.$$

Typically, $N(n, d)/N(k, d)$ is not an integer, so in this case the system which determines the coefficients c of the functions of k variables, say $Bc = p$, is underdetermined. Then we say that the chosen points form a basis if B has full rank. For all monomials we use the basis $x^{\mathbf{j}}/\sqrt{\mathbf{j}!}$. The scaling condition $\|a_j\|_2 = 1$ used for functions of 1 variable is replaced by the condition that the basis of each isotropic subspace be orthogonal, and we again choose the solution c of minimum norm. Since $\mathrm{Sp}(n) \subset \mathrm{O}(n)$, the condition number is invariant under symplectic (hence orthogonal) linear maps.

For $n = 6$ and $d = 3$, for example, the conjecture implies the 56 monomials can be expressed as a sum of $s^* = 56$, $\lceil 56/4 \rceil = 14$, and $\lceil 56/10 \rceil = 6$ cubics in $k = 1, 2$, and 3 variables, respectively. In the three cases we have to choose s^* points in a manifold of dimension 5, 7, and 6 respectively.

An isotropic subspace with $k = \tilde{n} := n/2$ is called *Lagrangian*, and $\mathrm{IG}(n, \tilde{n}) =: \mathrm{LG}(n)$, the *Lagrangian Grassmannian*, of dimension $\frac{1}{2}\tilde{n}(\tilde{n} + 1)$ [14]. The homeomorphism $\mathrm{LG}(n) \cong \mathrm{U}(\tilde{n})/\mathrm{O}(\tilde{n})$

13

can be realized concretely in the following way: let the subspace consist of the functions $F_i = a_i^{\mathrm{T}} x$ for $a_i \in \mathbb{R}^n$, and let the ith row of $A \in \mathbb{R}^{\tilde{n} \times n}$ be a_i^{T}. Let $A = [B|C]$, where $B, C \in \mathbb{R}^{\tilde{n} \times \tilde{n}}$. Then A specifies a Lagrangian subspace iff the rows of A are linearly independent and BC^{T} is symmetric. Now consider the complex matrix $U = B + \sqrt{-1}C$. Since $BC^{\mathrm{T}} - B^{\mathrm{T}}C = 0$, $UU^\dagger$ is real. Then by taking linear combinations of the rows of A (i.e., by a change of basis of the subspace), we can make $UU^\dagger = I$, that is, $U \in \mathrm{U}(\tilde{n})$. However, the unitary matrices U and QU represent the same subspace for any orthogonal $Q \in \mathrm{O}(\tilde{n})$; we thus identify them to get $\mathrm{LG}(n) \cong \mathrm{U}(\tilde{n})/\mathrm{O}(\tilde{n})$. See [4] for a description of $\mathrm{LG}(4)$ as a nontrivial S^2 bundle over S^1. In practice we mapped $\mathrm{U}(n) \to \mathrm{LG}(n)$ as above and represented points in $\mathrm{LG}(n)$ by unitary matrices, even though this map is not 1–1.

The number of pieces now required is quite small and, importantly, it is bounded with respect to n.

Proposition 3 $s^* := \lceil N(n,d)/N(n/2,d) \rceil \leq 2^d$ *for all* n.

Proof Let $n = 2m$. We have

$$
\begin{aligned}
\frac{N(2m,d)}{N(m,d)} &= \frac{\binom{2m+d-1}{d}}{\binom{m+d-1}{d}} \\
&= \frac{m!(2m+d-1)!}{(2m)!(m+d-1)!} \\
&= \frac{(2m)\ldots(2m+d-1)}{(m)\ldots(m+d-1)} \\
&\leq 2^d \\
&= 2^d \left(1 - \frac{(d-1)(d-2)}{4m} + \mathcal{O}(m^{-2}) \right).
\end{aligned}
$$

$\bullet$

Thus we expect to be able to express cubics in any number of variables, for example, as a sum of at most 8 integrable pieces. The number of pieces conjectured is given in Table 2.

In the case $n = 6$, Rangarajan [18] has suggested using a set of points in $\mathrm{LG}(n)$ on which a discrete subgroup of $\mathrm{Sp}(n)$ acts transitively (as the icosahedron group acts on the verticies of the icosahedron in $\mathbb{R}^3$). Based on the data for packing points in Grassmannians in [3, 21], it seems likely that such configurations will be optimal for certain choices of n and d. The drawbacks are that (i) the discrete subgroups of the classical groups are not known in general, (ii) such a subgroup can act transitively only on a very restricted set of numbers of points s, and (iii) such configurations may be singular. In our experiments, singularity depends on the symmetry group and on the orbit decomposition of the set of points.

We have not found any candidate good sets of s^* points parameterized by n or d. Although s^* is bounded as $n \to \infty$, the number $N(n,d)$ of constraints grows rapidly, so that it may not be possible to achieve uniformly good conditioning. For example, it is necessary that the independent equations corresponding to all choices of d out of n variables be well-conditioned.

	n									
d	2	4	6	8	10	12	14	16	18	20
2	3	4	4	4	4	4	4	4	4	4
3	4	5	6	6	7	7	7	7	7	7
4	5	7	9	10	11	11	12	12	13	13
5	6	10	12	15	16	18	19	20	21	22
6	7	12	17	21	24	27	30	32	34	36
7	8	15	22	29	35	41	46	50	54	58
8	9	19	29	39	50	59	68	77	85	92

Table 2: The number of pieces required to split a Hamiltonian of degree d in n variables, $2 \leq n \leq 20$, into a sum of functions of $n/2$ fixed, commuting, linear variables.

We close this section by posing an interesting problem in symplectic geometry. The minimum s required to split a generic polynomial into a sum of functions of one variable is known as the (algebraic) Waring problem [1, 12]. We thus have the

Symplectic Waring problem. For each d, each (even) n, and each $1 \leq k \leq n/2$, determine the minimum s required to split a generic polynomial of degree d in n variables into a sum of polynomials of k Poisson-commuting linear functions.

A direct counting of parameters suggests that

$$ s = \left\lceil \frac{N(n,d)}{N(k,d) + \dim \mathrm{IG}(n,k)} \right\rceil . \tag{6} $$

We verified (6) numerically for some values of n and d, namely $n = 4$ and $d \leq 10$; $n = 6$ and $d \leq 4$; $n = 8$ and $d \leq 3$; and $n = 10$, $d = 3$. That is, in these 15 cases we did not discover any obstructions in the symplectic Waring problem as exist in four cases of the standard Waring problem. On the other hand, a related problem, of expressing polynomials in 3 variables as sums of polynomials in 2 variables, has been studied [5] and typically *does* exhibit obstructions.

Note that for $k = n/2$ and $d = 2$, we get $s = 2$. Can every quadratic in n variables be written as a sum of two shears? Can this be done explicitly?

4.1 Diagonal basis

Instead of choosing s *arbitrary* points on $\mathrm{LG}(n)$, we now consider restricting the points so as to get an easier optimization problem and (conceptually) simpler decompositions. We look for submanifolds of $\mathrm{LG}(n)$ which are geometrically simple and have bases with a lot of zeros in them, which makes the associated vector fields faster to integrate.

First, consider the submanifold of $\mathrm{LG}(n)$ given by the Lagrangian subspaces

$$ \mathrm{span}(a_1 q_1 + b_1 p_1, \ldots, a_{n/2} q_{n/2} + b_{n/2} p_{n/2}) \tag{7} $$

d	n 2	4	6	8	10	12	14	16	18	20
2	3	4	4	4	4	4	4	4	4	4
3	4	6	8	8	8	8	8	8	8	8
4	5	9	12	16	16	16	16	16	16	16
5	6	12	18	24	32	32	32	32	32	32
6	7	16	27	36	48	64	64	64	64	64
7	8	20	36	54	72	96	128	128	128	128
8	9	25	48	81	108	144	192	256	256	256

Table 3: As Table 2, but with the ith variable a combination of q_i and p_i only (a 'diagonal' basis).

where (a_i, b_i) specifies a line through the origin in $\mathbb{R}^2$. That is, the ith variable depends on q_i and p_i only. These subspaces are specified by coordinates $a_i = \cos\theta_i$, $b_i = \sin\theta_i$, with θ, $-\theta$ identified. This submanifold is isomorphic to $\mathbb{T}^{n/2}$.

Splitting using this submanifold was introduced by Irwin [13], who considered $n = 4$ and $n = 6$, but did not consider the need to optimize the basis, or that a given configuration in $\mathbb{T}^{n/2}$ might be singular.

Second, consider the isotropic subspaces

$$\mathrm{span}(aq_2 + bp_1, aq_1 + bp_2)$$

where (a, b) specifies a line through the origin in $\mathbb{R}^2$. This describes a circle inside $\mathrm{LG}(n)$. To combine this and the previous class of subspaces, we can choose any set of $k \leq n/4$ pairs of indices from $\{1, \dots, n/2\}$; to each pair (i, j) we attach the circle $(aq_i + bp_j, aq_j + bp_i)$, and to each of the remaining singletons l we attach the circle $aq_l + bp_l$. Each such set of indices gives a torus $\mathbb{T}^{n/2-k} \subset \mathrm{LG}(n)$. There are $m(m-1)/2$ choices for the first pair, where $m = n/2$, $(m-2)(m-3)/2$ choices for the second pair, and so on; but the order of the pairs doesn't matter, giving

$$\frac{m \dots (m - 2k + 1)}{2^k k!}$$

such pairs. For example, for $n = 8$ ($\dim \mathrm{LG}(8) = 10$) we can choose 0 pairs (corresponding to $\mathbb{T}^4 \subset \mathrm{LG}(8)$); 1 pair in 12 ways (giving 12 copies of $\mathbb{T}^3 \subset \mathrm{LG}(8)$); or 2 pairs in 3 ways (giving 3 copies of $\mathbb{T}^3 \subset \mathrm{LG}(8)$). The total of s points can be distributed amongst these submanifolds in any way, subject to nonsingularity of the configuration. In the example below ($n = 4$, $d = 3$, $s = 5$), we find that putting either 2, 3, or 4 points on $\mathbb{T}^2$ (0 pairs) and the rest on $\mathbb{T}^1$ (1 pair) does give nonsingular, and in fact quite good, configurations.

The most extreme case is to choose 0 pairs and put all s points in $\mathbb{T}^{n/2}$. We call such a configuration a 'diagonal' basis. This does, however, necessitate increasing the value of s, although without altering the asymptotics of s for large n and d too severely.

Proposition 4 *Let $m = n/2$, $p = \lceil d/m \rceil$ and $i = mp - d$. Then there are diagonal bases with $s := p^i(p+1)^{m-i}$ points. We have $s < 2^d$ for all $n < 2d$ and $s = 2^d$ for all $n \geq 2d$.*

16

Proof   Consider the monomials of degree d which have degree d_i in (q_i, p_i), so that $\sum_{i=1}^{m} d_i = d$. Such terms in P are to be matched with terms in the assumed expansion of the form

$$P_j(a_1 q_1 + b_1 p_1, \ldots, a_m q_m + b_m p_m),$$

where P_j is a polynomial of degree d_i in its ith argument. All such monomials occur in P_i, so there are $\prod_{i=1}^{m}(d_i + 1)$ such monomials. The equations which match *all* monomials in P break into independent sets, one for each partition of d into the d_i. So the configuration must be chosen so that each such set is nonsingular. The partition with $d_1 = \ldots = d_i = p - 1$, $d_{i+1} = \ldots = d_m = p$, with i, p as in the theorem, gives the largest set of monomials, namely, $s = p^i(p+1)^{m-i}$. For, suppose there is a partition with larger $\prod_{i=1}^{m}(d_i + 1)$. Let its smallest degree be $d_1 = p - 1 - a$ and its largest be $p + b$ where necessarily a, $b \geq 0$, with a, b not both equal to 0. But then the number of monomials can be increased by increasing d_1 by 1 and decreasing d_m by 1, for

$$(p - a + 1)(p + b) = p^2 + (b + 1 - a)p - ab + b \geq (p - a)(p + b + 1) = p^2 + (b + 1 - a)p - ab - a.$$

Thus this partition of d did not give a maximum number of monomials and we have a contradiction.

The equations corresponding to each partition of d are of Vandermonde type, with the entry corresponding to a given monomial in (q, p) being the same monomial in (a, b). They generically have full rank. We have shown that the value of s given in the theorem ensures that the largest such set of equations has as many columns as rows. Hence this value of s makes generic configurations nonsingular.

●

As $d \to \infty$ for fixed n, $s = \mathcal{O}(d^{n/2})$, as in the case of a general basis in $\mathrm{LG}(n)$, although the constant is larger. The values of s are given in Table 3.

4.2   Cost of the integrators

How do the different types of integrators compare for cost? We compare splitting into functions of 1 variable (many needed, but fast to evaluate) vs. functions of $n/2$ variables.

Consider evaluating monomials by multiplying them out in the form $x_i x_j \ldots$, which costs (ignoring single operations) d operations. The vector field $J\nabla H$ has n components, each a polynomial of degree $d - 1$, so costs about $2ndN(n, d-1)$ to evaluate.

Integrating by splitting into functions of 1 variable costs $2n$ operations to evaluate $a^T x$, $2n$ to update x, and d operations to evaluate the polynomials of 2 variable, or $4n + d$ per piece, for a total of $N(n, d)(4n + d)$. The cost factor compared to simply evaluating the vector field is $(n + d)(4n + d)/(2nd^2) = \mathcal{O}(n/d^2)$ as $n \to \infty$. We have lost a power of n because the cost benefit in dropping from degree d (in the Hamiltonian) to degree $d - 1$ (in the vector field) is much less for polynomials in 1 variable than for polynomials in n variables. Still, there is a fairly big range of n values for which the cost is not prohibitive.

For functions of $n/2$ variables, we have to integrate $N(n, d)/N(n/2, d)$ vector fields, each with $n/2$ components of degree $d - 1$. The cost factor compared to evaluating the vector field is therefore approximately

$$\frac{1}{2} \frac{N(n, d) N(n/2, d-1)}{N(n/2, d) N(n, d-1)} = \frac{n + d}{n + 2d}$$

—that is, it is actually faster to integrate than to evaluate the vector field! This is possible because the splitting has taken into account the fact that the vector field is Hamiltonian, but (our) evaluation has not.

5   Example: Cubics in 4 variables.

We now consider the case $d = 3$, $n = 4$ in detail. There are $N(4, 3) = 20$ monomials. We have considered four approaches to splitting these polynomials.

1. Choosing 20 points in $\mathbb{P}^3$, i.e. using functions of 1 variable. We performed a local minimization of the condition number κ, where the initial guess was Sloane's optimal packing of 20 lines in $\mathbb{R}^4$ [21]. This gave a solution with $\kappa = 0.499$; we did not find a better solution from other starting guesses. The solution has a lot of symmetry—its singular values occur in sets with multiplicities 4, 4, 4, and 8 (see Figure 3).

2. Choosing 20 points in $\mathbb{P}^3$ with full permutation symmetry. The basis is $(1, 0, 0, 0)$ and its 4 permutations; $(1, \pm 1, 0, 0)/\sqrt{2}$ and their 8 permutations; and $(1, 1, 1, 0)/\sqrt{3}$ and its 4 permutations. This basis has $\kappa = 0.707$, which is quite good considering no global optimization has been done.

3. Choosing 6 points in $\mathbb{T}^2$, i.e., functions of 2 variables, but restricting to diagonal bases. The solution with smallest κ that we have found has angles in point j (see Eq. (7)) given by $\theta_1 = \Theta_{1,j}$, $\theta_2 = \Theta_{2,j}$ where

$$\Theta = \pi \begin{bmatrix} x & -x & \frac{1}{2} - x & x - \frac{1}{2} & \frac{1}{2} & 0 \\ x & -x & x - \frac{1}{2} & \frac{1}{2} - x & 0 & \frac{1}{2} \end{bmatrix} \tag{8}$$

where $x = 0.1568$. The 6 points in $\mathbb{T}^2$ are shown in Figure 2. This configuration has $\kappa = 0.481$. It has a symmetry group of order 8, generated by the involutions (in their action on the Lagrangian subspaces) $p \mapsto -p$, $(q_1, q_2, p_1, p_2) \mapsto (q_2, q_1, p_2, p_1)$, and $(q_1, q_2, p_1, p_2) \mapsto (p_1, -p_2, q_1, q_2)$. The 6 points are shown in Figure 2.

4. Choosing 5 points in LG(4). Even though LG(4) is only 3 dimensional, giving a 15 dimensional optimization problem, a naive approach did not succeed in solving it directly. Instead, we restricted the search to the submanifolds described in Section 4.1; we also studied the symmetry of the previous solution and imposed part of that symmetry on the solution. For $n = 2$ the submanifolds are $\mathbb{T}^2$ as in the previous case, and the circle defined by $F_1 = \cos y q_2 + \sin y p_1$, $F_2 = \cos y q_1 + \sin y p_2$. These latter points automatically have the symmetry $(q_1, q_2, p_1, p_2) \mapsto (q_2, q_1, p_2, p_1)$.

   We now consider various distributions of the 5 points between the torus and the circle. With 4 points on the torus, it turns out that the orbit structure under the symmetry is crucial:
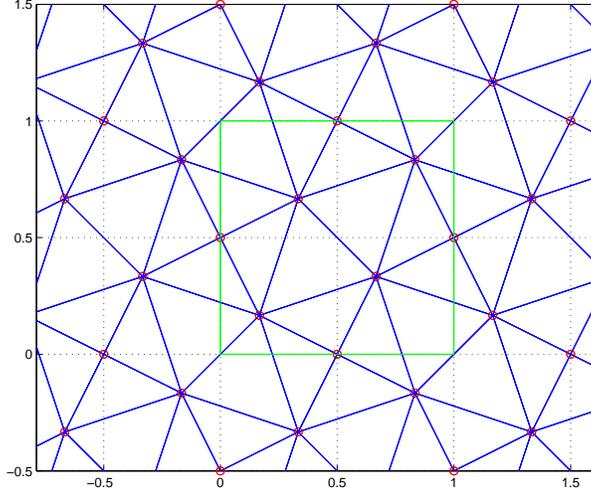
Figure 2: The best set of 6 points in $\mathbb{T}^2$ (see Eq. (8)). The torus is the unit square with periodic boundaries.

orbit types $1+1+1+1$ and $2+2$ turn out not to yield bases, i.e. B is always singular. Only orbit type $1+1+2$, two fixed points and a point of period 2, yields a basis. The best such has $\kappa = 0.537$. The 4 points on the torus are symmetric under $p \rightarrow -p$, but not the point on the circle, suggesting that even partial symmetries are worthwhile.

With 3 points on the torus, with orbit type $1+2$, the best basis has $\kappa = 0.474$. Now the 2 points on the circle are symmetric under $p \rightarrow -p$, but the 3 points on the torus are not. This is the best set of 5 points in $LG(4)$ that we have found. It corresponds to the basis

$$
(c_1 q_1 + s_1 p_1, c_1 q_2 + s_1 p_2)
$$
$$
(c_2 q_1 + s_2 p_1, c_3 q_2 + s_3 p_2)
$$
$$
(c_3 q_1 + s_3 p_1, c_2 q_2 + s_2 p_2)
$$
$$
(c_4 q_2 + s_4 p_1, c_4 q_1 + s_4 p_2)
$$
$$
(c_4 q_2 - s_4 p_1, c_4 q_1 - s_4 p_2)
$$

with $c_i = \cos \pi x_i$, $s_i = \sin \pi x_i$, and angles $x_1 = 0.6485$, $x_2 = 0.2256$, $x_3 = 0.5200$, and $x_4 = 0.1339$.

With 2 points on the torus, of period 2, the best basis has $\kappa = 0.537$.

The sets of 20 singular values are shown for each of the 4 solutions in Figure 3.

We then did a simple trial which involved choosing 200 random initial conditions on $\|x\| = 1$ and 200 random polynomials on $\|p\| = 1$, and calculating various measures of the local truncation errors of the energy for each basis. The results are shown in Table 4. The range in errors was only a factor of 1.4, with basis 1 the best, followed by 2, 4, and 3. The hierarchical basis (number 2) performs extremely well. Although basis 4 has the smallest condition number, it also has more terms in its local error (naively) from the triple Poisson brackets of the basis vectors. A similar effect penalizes basis 4; a general set of 5 points in $LG(4)$, without so many zeros, might suffer even more. Finally, the methods all have similar costs; the particular bases taken here cost 360 operations per step for bases 1–3, and 300 for basis 4.
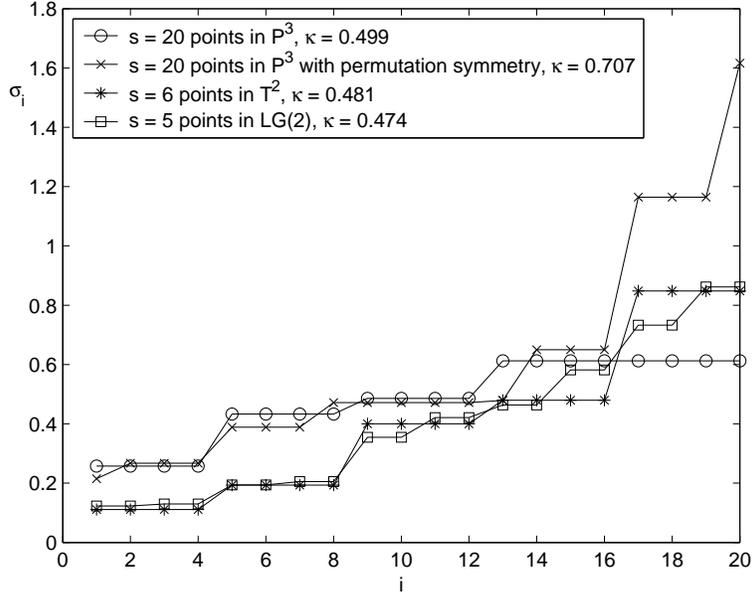
19

Figure 3: The 20 singular values σ_i of each of 4 different bases compared. Each basis allows one to split a cubic Hamiltonian in 4 variables into a sum of shears, where coordinate i of the Hamiltonian (in a suitable basis) is amplified by a factor σ_i.

So in this, simplest, case, all approaches are roughly competitive, and the clear winner for simplicity is the hierarchical basis 2.

6   Volume-preserving systems

In this section we consider ODEs of the form $\dot{x} = f(x)$ where $\nabla \cdot f = 0$, i.e., whose flow preserves Euclidean volume. Again, we restrict to polynomial f of degree d. We consider splitting a degree d, divergence-free vector field on $\mathbb{R}^n$ into a sum of s shears depending on s fixed sets of k variables, $1 \leq k \leq n-1$. How many pieces are required? There are $N(n,d)n$ independent vector fields of degree d. The divergence of such a vector field is a polynomial of degree $d-1$, which must

Case	s	κ	$\min E_i$	$\|E\|_{\text{geometric}}$	$\|E\|_2$	$\max E_i$	# ops
1	20	0.499	0.9	5.0	8.3	40	360
2	20	0.707	2.1	6.4	9.0	35	360
3	6	0.481	1.0	6.6	12.5	62	360
4	5	0.474	1.3	6.1	9.6	36	300

Table 4: Comparison of 4 different bases for splitting cubic Hamiltonians in 4 variables. The mean square truncation error e_i is calculated for each of 200 random polynomials, and 4 summaries of the range of $E_i := (100/(\Delta t)^3)e_i$ values reported. The last column is the number of operations to integrate the Hamiltonian for one time step at first order.

20

be identically zero. So there are $N(n,d)n - N(n, d-1)$ independent divergence-free vector fields. Each function in the splitting contains $n - k$ polynomials of degree d in k variables, hence each contributes $(n-k)N(k,d)$ unknowns. Then at least

$$
\begin{aligned}
r &:= \frac{N(n,d)n - N(n, d-1)}{(n-k)N(n-1,d)} \\
&= \frac{\frac{(n+d-1)!n}{d!(n-1)!} - \frac{(n+d-2)!}{(d-1)!(n-1)!}}{(n-k)\frac{(k+d-1)!}{d!(k-1)!}} \\
&= \frac{(n+d-2)!(n+d)(k-1)!}{(k+d-1)!(n-k)(n-2)!}
\end{aligned}
$$

pieces are required. Interestingly,

1. Splittings with the smallest number of pieces are not necessarily obtained by choosing $k = n - 1$. If $d < n$ then a smaller k minimizes r. For example, for $n = 6$ and $d = 2$, $k = 3$ gives splittings into 6 pieces, where $k = 5$ requires 8 pieces.

2. For $k = n - 1$, r is an integer: $r = n + d$. This suggests that there is something natural about this choice of k. Splittings in such a basis are then unique. Note that the cost of requiring the pieces to be shears is extremely small, considering that the original vector field depends on $n - 1$ scalar functions to begin with.

A basis is defined by a set of s points in $\mathrm{G}(n, k)$, the manifold of k-planes in $\mathbb{R}^n$. In coordinates, we choose an orthonormal basis $a_1, \dots, a_k$ for the k-plane, and complete it to an orthonormal basis of $\mathbb{R}^n$; the split vector fields then have the form

$$
\dot{x} = \sum_{i=1}^{n-k} f_i(a_1^{\mathrm{T}} x, \dots, a_k^{\mathrm{T}} x) a_{k+i}
$$

where the f_i are scalar functions.

Conjecture 2 *Generic sets of $\lceil r \rceil$ points in $\mathrm{G}(n, k)$ generate nonsingular configurations, i.e. bases. In the case $k = n - 1$, n of the $n + d$ required points can be taken as the n elementary unit vectors, which eliminates all monomials containing fewer than n variables.*

Some trials for small values of n and d support the conjecture. We anticipate that good packings of Grassmannians [21] will yield reasonably good bases and good initial guesses for optimization.

7 Conclusions

This splitting problem in the Hamiltonian setting has been studied in the particle accelerator community for some time, in the context of finding symplectic approximations of nearly symplectic

maps. The basic approach was introduced by Dragt and Neri in 1988 [8], while the diagonal basis was introduced by Irwin [13] and the LG(n) basis by Dragt and Abell [7] and studied further by Rangarajan [18]. Only [7] considers the importance of the conditioning of the basis; some of the sets of points suggested by [13] are in fact singular. These authors emphasize the cases $n = 4$ and $n = 6$ as most relevant to accelerators. The hierarchical basis was introduced by Scully [20].

The framework we have presented in this paper puts this problem in the context of several others: splitting in IG(n, k) for $1 \leq k \leq n/2$; non-canonical and Lie-Poisson systems; and volume preserving splitting in G(n, k) for $1 \leq k < n$. Its structure is revealed by considering all n and d; by its close connection with the problem of optimal experimental design in statistics; and with arranging points on spheres and homogeneous spaces. So far, all of the approaches we have considered are competitive, suggesting a thorough study of the associated optimization problems as an urgent next step. We leave this and the many natural extensions of this framework (for example, to general (non-structure-preserving) systems, to systems with a given set of linear symmetries and/or reversing symmetries, and to structure-preserving lattice maps) to the future.

References

[1] J. Alexander and A. Hirschowitz, Polynomial interpolation in several variables, *J. Algebraic Geometry* **4** (1995), 201–222.

[2] S. Blanes, Symplectic maps for approximating polynomial Hamiltonian systems, *Phys. Rev. E.* **65** (2002), 56703.

[3] J. H. Conway, R. H. Hardin and N. J. A. Sloane, Packing lines, planes, etc., packings in Grassmannian spaces, *Exp. Math.,* **5** (1996), 139-159.

[4] M. Dedo, and R. Giachetti, The Lagrangian Grassmannian $L(2)$, *Rend. Sem. Fac. Sci. Univ. Cagliari* **60** (1990), 65–89.

[5] E. Carlini, Geometric aspects of some polynomial decompositions, Ph.D. Thesis, Università di Pavia, 2003.

[6] A. J. Dragt, Lectures on Nonlinear Dynamics and Lie Methods with Applications to Accelerator Physics, preprint, 1993.

[7] A. J. Dragt and D. T. Abell, Symplectic maps and computation of orbits in particle accelerators, *Fields Inst. Comm.* **10**, 59–85, 1996.

[8] A. J. Dragt, F. Neri, G. Rangarajan, D. R. Douglas, L. M. Healy and R. D. Ryne, Lie algebraic treatment of linear and nonlinear beam dynamics, in *Annual Review of Nuclear and Particle Science* **38** (1988), 455–496.

[9] K. Feng and D.–L. Wang (1998), Variations on a theme by Euler, *J. Comput. Math.* **16**(2), 97–106.

[10] R. H. Hardin and N. J. A. Sloane, A new approach to the construction of optimal designs, *J. Stat. Planning & Inference* **37** (1993), 339–369.

[11] R. H. Hardin and N. J. A. Sloane, Codes (spherical) and designs (experimental), in *Different aspects of coding theory (San Francisco, CA, 1995)*, 179–206, Proc. Sympos. Appl. Math., 50, Amer. Math. Soc., Providence, RI, 1995.

[12] A. Iarrobino and V. Kanev, *Power Sums, Gorenstein Algebras and Determinantal Loci*, Springer Lecture Notes in Mathematics No. 1721, Springer, Berlin, 1999.

[13] J. Irwin, A multi-kick factorization algorithm for nonlinear maps, report SSC–228, 1989.

[14] P. Libermann and C.-M. Marle, *Symplectic Geometry and Analytical Mechanics*, D. Reidel, Dordrecht, 1987.

[15] R. I. McLachlan, Explicit Lie-Poisson integration and the Euler equations, *Phys. Rev. Lett.* **71** (1993), 3043–3046.

[16] R. I. McLachlan and G. R. W. Quispel, Splitting methods, *Acta Numer.* **11** (2002), 341–434.

[17] M. Mikosz, On classification of the linear Lagrangian and isotropic subspaces, *Demonstratio Math.* **30** (1997), 437–450.

[18] G. Rangarajan, Symplectic completion of symplectic jets, *J. Math. Phys.* **37** (1996), 4514–4542.

[19] E. B. Saff and A. B. J. Kuijlaars, Distributing many points on a sphere. *Math. Intelligencer* **19** (1997), 5–11.

[20] J. E. Scully, A search for improved numerical integration methods using rooted trees and splitting, Thesis, La Trobe University, 2002.

[21] N. J. A. Sloane, Packings in Grassmannian spaces,
`http://www.research.att.com/~njas/grass/index.html`