

Explicit volume-preserving splitting methods for divergence-free ODEs by tensor-product basis decompositions

ANTONELLA ZANNA

Matematisk institutt, Universitet i Bergen, P.O. Box 7800, N-5020 Bergen, Norway
Antonella.Zanna@math.uib.no

[Received on 11 September 2012; revised on 15 August 2013]

We discuss the construction of volume-preserving splitting methods based on a tensor product of single-variable basis functions. The vector field is decomposed as the sum of elementary divergence-free vector fields (EDFVFs), each of them corresponding to a basis function. The theory is a generalization of the monomial basis approach introduced in Xue & Zanna (2013, *BIT Numer. Math.*, **53**, 265–281) and has the trigonometric splitting of Quispel & McLaren (2003, *J. Comp. Phys.*, **186**, 308–316) and the splitting in shears of McLachlan & Quispel (2004, *BIT*, **44**, 515–538) as special cases. We introduce the concept of diagonalizable EDFVFs and identify the solvable ones as those corresponding to the monomial basis and the exponential basis. In addition to giving a unifying view of some types of volume-preserving splitting methods already known in the literature, the present approach allows us to give a closed-form solution also to other types of vector fields that could not be treated before, namely those corresponding to the mixed tensor product of monomial and exponential (including trigonometric) basis functions.

Keywords: geometric integration; volume preservation; splitting methods; elementary divergence-free vector fields.

1. Introduction

Volume preservation is an important property shared by several dynamical systems. The principal example is the velocity field for an incompressible fluid flow and applications include the study of turbulence and of mixing of flows (see Arnold & Khesin, 1998). The preservation of phase-space volume is also a key ingredient in methods that involve stochastics like the hybrid Monte Carlo algorithm (see Neal, 2011). Thus, preservation of volume by a numerical method for differential equations is a desirable property either in conjunction with other properties, like energy preservation or time reversal symmetry, or on its own. The latter case is the subject of this paper.

Designing volume-preserving numerical integrators is a hard task, especially given the existence of no-go theorems within the class of ‘standard’ methods: in Qin & Zhu (1993) it was proved that no standard numerical method, for example a Runge–Kutta scheme, is volume preserving for all linear vector fields. A similar result was proved in Kang & Shang (1995). Further, Iserles *et al.* (2007) and Chartier & Murua (2007) proved that no B-series method can be volume preserving for all possible divergence-free vector fields. The only methods in this class that can be volume preserving are the composition of exact flows of divergence-free subsystems. Thus, the design of efficient methods that preserve volume is still an open problem in geometric integration (see McLachlan & Scovel, 1998).

Despite the existence of no-go theorems, volume-preserving numerical methods can be constructed by other techniques, for instance using generating functions (see e.g. Scovel, 1991; Shang, 1994; Quispel, 1995; Xue & Zanna, 2014) or by splitting methods. One of the earliest volume-preserving splitting methods is the splitting method by Kang & Shang (1995), decomposing the vector field into the sum

of essentially two-dimensional Hamiltonian fields, which are then solved by a (typically implicit) symplectic method.

Because of the difficulty of addressing the general space of divergence-free vector fields, recent efforts have concentrated on smaller, yet still interesting, function spaces, for instance, the space of polynomial fields. An earlier paper on the splitting of polynomial vector fields is [McLachlan & Quispel \(2004\)](#), dealing mainly with the Hamiltonian case. In [Xue & Zanna \(2013\)](#), we presented a novel approach for arbitrary polynomial divergence-free vector fields, by expanding the divergence equation in terms of the monomial basis, that allowed us to develop explicit volume-preserving splitting methods. It turns out that the approach can be generalized to an arbitrary tensor basis framework, which embraces several known splitting methods as special cases, namely the monomial splitting of [Xue & Zanna \(2013\)](#), the splitting in shears of [McLachlan & Quispel \(2004\)](#) and the trigonometric splitting of [Quispel & McLaren \(2003\)](#). The main idea is simple: to expand the divergence equation in terms of a tensor product of one-dimensional basis functions. Each basis function gives rise to a uniquely determined (up to an *off-diagonal contribution*) elementary divergence-free vector field (EDFVF). Away from some *critical points*, these EDFVFs are proved to be integrable. Further, we identify solvable EDFVFs, recovering the monomial basis and the exponential basis (and the mixed tensor product of these). The resulting composition methods are thus explicit and volume preserving.

Finally, we also consider expansions in EDFVFs for divergence-free fields with measure $\mu(\mathbf{x}) \neq 1$, and show that the noncanonical case can be viewed as a time-reparametrization of the canonical one. Thus, the results and methods presented in this paper can be applied to the case $\nabla \cdot (\mu \mathbf{f}) = 0$.

2. The basis-decomposition approach

We consider the problem of solving the differential equation

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \quad (2.1)$$

subject to the divergence-free condition

$$\nabla \cdot \mathbf{f} = 0, \quad (2.2)$$

by splitting the vector field \mathbf{f} as the sum of EDFVFs that are integrable by analytical methods. As the exact flow of each of the elementary vector fields is itself volume preserving, the resulting composition method is volume preserving.

In [McLachlan *et al.* \(2008\)](#), it was proposed to decompose any divergence-free vector field $\mathbf{f}(\mathbf{x})$ into a *diagonal* and *off-diagonal* part,

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}^{\text{diag}}(\mathbf{x}) + \mathbf{f}^{\text{offdiag}}(\mathbf{x}),$$

where, componentwise, $f_i^{\text{diag}}(\mathbf{x})$ is the collection of terms in $f_i(\mathbf{x})$ that depend on x_i (i.e. $\partial_{x_i} f_i^{\text{diag}}(\mathbf{x}) \neq 0$), while $\mathbf{f}^{\text{offdiag}}(\mathbf{x})$ is, componentwise, the collection of terms in f_i that do not depend on x_i , i.e. $\partial_{x_i} f_i^{\text{offdiag}}(\mathbf{x}) = 0$. From the definition of divergence, it is clear that only the coefficients of the diagonal part are directly involved in the divergence-free condition (2.2); therefore, vector fields with zero diagonal part are automatically divergence-free. Further, there exists a generic volume-preserving splitting technique to solve for these off-diagonal vector fields, namely, the splitting into canonical shears: one integrates one variable at a time, keeping the remaining variables constant; see [McLachlan & Quispel \(2004\)](#) and [McLachlan *et al.* \(2008\)](#). Vector fields with nonzero diagonal are more difficult to treat; thus,

without loss of generality, we assume that the function \mathbf{f} has zero off-diagonal part, i.e.

$$\mathbf{f} = \mathbf{f}^{\text{diag}}(\mathbf{x}), \quad \mathbf{f}^{\text{offdiag}}(\mathbf{x}) = 0, \quad (2.3)$$

unless otherwise mentioned.

PROPOSITION 2.1 Let $\mathbf{f}(\mathbf{x})$ be a given vector field, with divergence $p(\mathbf{x}) = \nabla \cdot \mathbf{f}(\mathbf{x})$, and assume that the latter can be expanded in a set of basis functions $\{\phi_{\mathbf{j}}(\mathbf{x})\}_{\mathbf{j} \in \mathcal{J}}$,

$$p(\mathbf{x}) = \sum_{\mathbf{j} \in \mathcal{J}} p_{\mathbf{j}} \phi_{\mathbf{j}}(\mathbf{x}).$$

For each $\mathbf{j} \in \mathcal{J}$, there exists a uniquely determined vector field $\mathbf{F}_{\mathbf{j}}(\mathbf{x})$ such that

$$\nabla \cdot \mathbf{F}_{\mathbf{j}}(\mathbf{x}) = p_{\mathbf{j}} \phi_{\mathbf{j}}(\mathbf{x}). \quad (2.4)$$

If \mathbf{f} is divergence-free, then the differential equation

$$\dot{\mathbf{x}} = \mathbf{F}_{\mathbf{j}}(\mathbf{x}) \quad (2.5)$$

is also divergence-free and, moreover, the decomposition $\mathbf{f}(\mathbf{x}) = \sum_{\mathbf{j}} \mathbf{F}_{\mathbf{j}}(\mathbf{x})$ is unique.

Proof. Since the basis functions are independent,

$$\nabla \cdot \mathbf{f}(\mathbf{x}) = \sum_{\mathbf{j} \in \mathcal{J}} p_{\mathbf{j}} \phi_{\mathbf{j}}(\mathbf{x}) \implies p_{\mathbf{j}} = 0$$

for all indices $\mathbf{j} \in \mathcal{J}$ whenever \mathbf{f} obeys (2.2). Consider next vector fields $\mathbf{F}_{\mathbf{j}}$, constructed so that

$$(\mathbf{F}_{\mathbf{j}})_i = (\mathbf{a}_{\mathbf{j}})_i \int \phi_{\mathbf{j}}(\mathbf{x}) \, dx_i, \quad (2.6)$$

where $(\mathbf{a}_{\mathbf{j}})_i$ is the component of $\partial_i f_i(\mathbf{x})$ along $\phi_{\mathbf{j}}(\mathbf{x})$. Note that the $\mathbf{F}_{\mathbf{j}}$ s are uniquely defined up to a function $h_i(\mathbf{x})$ independent of x_i , which we can choose to be zero, as it would only contribute to the off-diagonal part. By construction, $\mathbf{F}_{\mathbf{j}}(\mathbf{x})$ is such that $\nabla \cdot \mathbf{F}_{\mathbf{j}}(\mathbf{x}) = p_{\mathbf{j}} \phi_{\mathbf{j}}(\mathbf{x}) = 0$ since $p_{\mathbf{j}} = 0$. Hence, $\mathbf{F}_{\mathbf{j}}$ is divergence-free. Finally, by collecting all terms in \mathbf{f} by picking out those which contribute to each $p_{\mathbf{j}}$, we are able to split \mathbf{f} as the sum of EDFVs, $\mathbf{f} = \sum_{\mathbf{j} \in \mathcal{J}} \mathbf{F}_{\mathbf{j}}$. \square

A practical way to find the components of the vectors $\mathbf{a}_{\mathbf{j}}$ in (2.6) is the following: let $\mathcal{J}^{[0]} = \{\}$ be the initial (empty) list of indices for the basis functions. Starting with $i = 1$, for $i = 1, \dots, n$, expand $\partial_i f_i$ in the given basis. If new basis functions are added to the expansion, the corresponding indices are added to the list of indices $\mathcal{J}^{[i]}$. For each index $\mathbf{j} \in \mathcal{J}^{[i]}$, the coefficient $(\mathbf{a}_{\mathbf{j}})_i$ is determined. When $i = n$, $\mathcal{J}^{[n]} = \mathcal{J}$ is the final list of indices of the expansion and the $\mathbf{a}_{\mathbf{j}}$ are completely determined.

It might happen that some of the $\mathcal{J}^{[i]}$ s, hence \mathcal{J} , are infinite. This is the case when some of the $\partial_i f_i$ s expand in infinite terms. In practice, it is desirable to work with a finite number of basis elements: we assume that each $\partial_i f_i$ can be projected to the same finite-dimensional space spanned by a finite number of basis functions. Correspondingly, we obtain a projection of the vector field \mathbf{f} onto a finite-dimensional set of divergence-free vector fields. The dimension of such a set is ideally chosen by the problem owner not to alter the dynamic of the system of interest in a significant manner. Alternatively, we assume

that the remainder of the finite-dimensional projection can be treated by some other volume-preserving numerical method. Thus, without further ado, we will assume \mathcal{J} to be finite.

DEFINITION 2.2 Given the basis $\{\phi_{\mathbf{j}}(\mathbf{x})\}_{\mathbf{j} \in \mathcal{J}}$, the vector field $\mathbf{F}_{\mathbf{j}}(\mathbf{x})$ obeying (2.4) will be called an EDFVF associated to the basis function $\phi_{\mathbf{j}}(\mathbf{x})$.

The main idea of the paper can be summarized as follows: decompose the *divergence function* using the basis functions $\{\phi_{\mathbf{j}}(\mathbf{x})\}_{\mathbf{j} \in \mathcal{J}}$,

$$p(\mathbf{x}) = \sum_{\mathbf{j} \in \mathcal{J}} p_{\mathbf{j}} \phi_{\mathbf{j}}(\mathbf{x}).$$

For each $\mathbf{j} \in \mathcal{J}$,

- identify the terms in $\mathbf{f}(\mathbf{x})$ that contribute to $p_{\mathbf{j}}$ to obtain $\mathbf{F}_{\mathbf{j}}(\mathbf{x})$ (EDFVF);
- integrate $\dot{\mathbf{x}} = \mathbf{F}_{\mathbf{j}}(\mathbf{x})$ exactly (if possible) or by some volume preserving method.

The composition resulting from the flows of the $\mathbf{F}_{\mathbf{j}}$ is then volume preserving by construction. The main goal is to find suitable basis functions, in the sense that the corresponding vector fields must be easy to integrate exactly.

3. Tensor-product bases

Consider the case when the basis functions are the *tensor product* of one-dimensional function bases,

$$\phi_{\mathbf{j}}(\mathbf{x}) = \phi_{j_1}(x_1) \cdots \phi_{j_n}(x_n), \quad (3.1)$$

$$j_l \in \mathcal{J}_l, \quad \mathbf{j} = (j_1, \dots, j_n) \in \mathcal{J} = \mathcal{J}_1 \times \cdots \times \mathcal{J}_n.$$

Observe that the choice

$$\phi_{j_l}(x_l) = x_l^{j_l}, \quad \phi_{\mathbf{j}}(\mathbf{x}) = \mathbf{x}^{\mathbf{j}} = x_1^{j_1} \cdots x_n^{j_n}$$

corresponds to the case of the *monomial* basis for polynomials, discussed in detail in [Xue & Zanna \(2013\)](#).

PROPOSITION 3.1 Let $\phi_{\mathbf{j}}(\mathbf{x})$ be in the form (3.1), for some arbitrary tensor-product basis. The EDFVF $\mathbf{F}_{\mathbf{j}}$ (2.5) corresponding to the multi-index $\mathbf{j} = (j_1, \dots, j_n)$ can be written as

$$\begin{aligned} \dot{x}_i &= (\mathbf{a}_{\mathbf{j}})_i \int \phi_{j_i}(x_i) dx_i \prod_{l=1/l \neq i}^n \phi_{j_l}(x_l) \\ &= (\mathbf{a}_{\mathbf{j}})_i \frac{\int \phi_{j_i}(x_i) dx_i}{\phi_{j_i}(x_i)} \phi_{\mathbf{j}}(\mathbf{x}), \quad i = 1, \dots, n, \end{aligned} \quad (3.2)$$

where the $\mathbf{a}_{\mathbf{j}} \in \mathbb{R}^n$ obey the condition

$$\sum_{i=1}^n (\mathbf{a}_{\mathbf{j}})_i = \mathbf{a}_{\mathbf{j}}^T \mathbf{1} = 0, \quad (3.3)$$

where $\mathbf{1} = (1, 1, \dots, 1)^T \in \mathbb{R}^n$.

Proof. The result is a corollary of Proposition 2.1, specific to the case of tensor bases. Note, in particular, that upon taking the divergence, the coefficients $(\mathbf{a}_j)_i$ obey the divergence-free condition $\mathbf{a}_j^T \mathbf{1} = p_j = 0$; thus (3.3) holds. \square

It is worth mentioning that the second line in (3.2) makes sense when $\phi_{j_i}(x_i) \neq 0$. This motivates the following definition.

DEFINITION 3.2 The points in which $\phi_{j_i}(x_i) = 0$ are called *critical points of the basis function* ϕ_{j_i} . Similarly, the points in which the $\int \phi_{j_i}(x_i) dx_i = 0$ are the *critical points of the primitives* of the basis functions. The collection of both is referred to as the *critical points* of the EDFVF \mathbf{F}_j , or simply *critical points*.

From now on, for ease of notation, we will discard the dependence of the vector \mathbf{a}_j on the index j , unless the index is needed for the sake of clarity.

THEOREM 3.3 (Integrability of EDFVFs of tensor bases) Consider the EDFVF (3.2) associated to the tensor basis function $\phi_j(\mathbf{x})$ ($\mathbf{a}_j \equiv \mathbf{a}$). Let $\mathbf{b}_1 = \mathbf{1}$ and $\mathbf{b}_2, \dots, \mathbf{b}_{n-1}$ be such that the set of vectors $\{\mathbf{a}, \mathbf{b}_1, \dots, \mathbf{b}_{n-1}\}$ are mutually orthogonal. Then, for any $l = 1, \dots, n-1$, the quantity $I_l = \prod_{i=1}^n (\int \phi_{j_i}(x_i) dx_i)^{(\mathbf{b}_l)_i}$ is an integral of the system, i.e.

$$I_l = \prod_{i=1}^n \left(\int \phi_{j_i}(x_i) dx_i \right)^{(\mathbf{b}_l)_i} = \text{const}, \quad l = 1, \dots, n-1, \quad (3.4)$$

for all t . In particular, away from the critical points, the system has $n-1$ independent integrals and is integrable.

Proof. Let \mathbf{b}_l be as above. Note that

$$\frac{d}{dt} \log \left(\int \phi_{j_i}(x_i) dx_i \right)^{(\mathbf{b}_l)_i} = (\mathbf{b}_l)_i \dot{x}_i \frac{\phi_{j_i}(x_i)}{\int \phi_{j_i}(x_i) dx_i}.$$

For each $i = 1, \dots, n$, multiply (3.2) by $(\mathbf{b}_l)_i \phi_{j_i}(x_i) / \int \phi_{j_i}(x_i) dx_i$ and sum over i :

$$\sum_{i=1}^n \dot{x}_i (\mathbf{b}_l)_i \frac{\phi_{j_i}(x_i)}{\int \phi_{j_i}(x_i) dx_i} = \sum_{i=1}^n a_i (\mathbf{b}_l)_i \phi_j(\mathbf{x}) = (\mathbf{a}^T \mathbf{b}_l) \phi_j(\mathbf{x}) = 0,$$

as a consequence of the orthogonality of the vectors. Hence, it follows that

$$\log \prod_{i=1}^n \left(\int \phi_{j_i}(x_i) dx_i \right)^{(\mathbf{b}_l)_i} = \text{const},$$

from which the first part of the statement follows by exponentiation of both sides of the equation. In passing, note that (3.4) holds also for any \mathbf{b} in the span of $\mathbf{b}_1 = \mathbf{1}, \mathbf{b}_2, \dots, \mathbf{b}_{n-1}$, i.e. for any \mathbf{b} orthogonal to \mathbf{a} .

As far as integrability is concerned, it is sufficient to show that the integrals are independent, i.e. that the linear combination

$$c_1 \nabla I_1 + c_2 \nabla I_2 + \cdots + c_{n-1} \nabla I_{n-1} = \mathbf{0} \quad (3.5)$$

admits only the trivial solution $c_1 = \cdots = c_{n-1} = 0$. For this purpose, observe that

$$(\nabla I_l)_i = \frac{I_l}{\int \phi_{j_i}(x_i) dx_i} \partial_{x_i} \left(\int \phi_{j_i}(x_i) dx_i \right)^{(\mathbf{b}_l)_i} = I_l (\mathbf{b}_l)_i \frac{\phi_{j_i}(x_i)}{\int \phi_{j_i}(x_i) dx_i}, \quad i = 1, \dots, n,$$

for any $l = 1, \dots, n-1$. Hence, the linear system (3.5) can be written in the form

$$\mathcal{E}[\mathbf{b}_1, \dots, \mathbf{b}_{n-1}] J \mathbf{c} = \mathbf{0},$$

where \mathcal{E} and J are the diagonal matrices with elements

$$(\mathcal{E})_{i,i} = \frac{\phi_{j_i}(x_i)}{\int \phi_{j_i}(x_i) dx_i}, \quad J_{i,i} = I_i$$

and dimensions $n \times n$ and $(n-1) \times (n-1)$, respectively. It is immediate to see that the matrix $[\mathbf{b}_1, \dots, \mathbf{b}_{n-1}]$ has full rank. The matrix J has full rank as long as each of the integrals I_i is nonzero, i.e. each of the quantities $\int \phi_{j_i}(x_i) dx_i \neq 0$. Similarly, the matrix \mathcal{E} has full rank, provided that $\phi_{j_i}(x_i) \neq 0$. Away from the critical points of the basis functions and their primitives, the system has full rank and the integrals are independent. Thus, the system is integrable. \square

DEFINITION 3.4 (Diagonalizable EDFVFs) Consider the EDFVF \mathbf{F}_j corresponding to the multi-index $\mathbf{j} = (j_1, \dots, j_n)$,

$$\dot{x}_i = a_i \frac{\int \phi_{j_i}(x_i) dx_i}{\phi_{j_i}(x_i)} \phi_j(\mathbf{x}), \quad i = 1, \dots, n, \quad \mathbf{a}^T \mathbf{1} = 0,$$

away from its critical points. We say that $\mathbf{F}_j(\mathbf{x})$ is *diagonalizable* if $\phi_j(\mathbf{x})$ has a closed-form expression.

Diagonalizability means that $\phi_j(\mathbf{x})$ can be treated as a function of t only. Thus, each x_i obeys a differential equation with variable coefficients, indicating that one might find classes of diagonalizable EDFVFs whose solution can be given in closed form. For this purpose, we study the behaviour of the basis function $\phi_j(\mathbf{x})$ as a function of time.

THEOREM 3.5 (Solvable diagonalizable EDFVFs) Consider the EDFVF \mathbf{F}_j away from its critical points. The tensor basis function $\phi_j(\mathbf{x}) = \phi_{j_1}(x_1) \cdots \phi_{j_n}(x_n)$ obeys the differential equation

$$\frac{d}{dt} \phi_j(\mathbf{x}) = \left(\sum_{i=1}^n a_i \frac{\phi'_{j_i}(x_i) \int \phi_{j_i}(x_i) dx_i}{\phi_{j_i}(x_i)^2} \right) \phi_j(\mathbf{x})^2. \quad (3.6)$$

Moreover, (3.6) is solvable if and only if the $\phi_{j_i}(x_i)$ are monomials (up to a shift) or exponentials (including Fourier basis functions).

Proof. Equation (3.6) is an immediate consequence of $\frac{d}{dt} \phi_{j_i}(x_i) = \phi'_{j_i}(x_i) \dot{x}_i$ and (3.2); hence the first part of the statement follows.

To prove the second part of the statement, we observe that (3.6) has the form

$$\frac{d}{dt}y = gy^2,$$

where $y(t) = \phi_j(\mathbf{x})(t)$. Clearly, the equation is closed if $g \equiv c$, a constant, or if $g \equiv g(t)$ is a function of time only. This latter case is equivalent to the case $g \equiv c$ upon the time reparametrization $d\tau = g(t) dt$.

We commence by analysing the constant case,

$$\frac{\phi'_{j_i}(x_i) \int \phi_{j_i}(x_i) dx_i}{\phi_{j_i}(x_i)^2} = c_i, \quad (3.7)$$

where $c_i \neq 0$ (the case $c_i = 0$ corresponds to constant basis functions and is uninteresting). In this case, the basis functions obey the differential equation

$$\frac{d}{dt}\phi_j(\mathbf{x}) = \left(\sum_{i=1}^n a_i c_i \right) \phi_j(\mathbf{x})^2, \quad (3.8)$$

which has solution

$$\phi_j(\mathbf{x})(t) = \frac{\phi_j(\mathbf{x})(t_0)}{1 - [\phi_j(\mathbf{x})(t_0) \sum_{i=1}^n a_i c_i] (t - t_0)} \quad (3.9)$$

in $[t_0, t]$ for $t - t_0$ sufficiently small.

Setting $\Phi(z) = \int \phi_{j_i}(z) dz$, the condition (3.7) is equivalent to Φ being a solution of the differential equation

$$\Phi'' \Phi = c_i (\Phi')^2, \quad (3.10)$$

where the dash denotes differentiation with respect to the variable z . Dividing both sides by $\Phi \Phi'$ and integrating, we obtain

$$\log \Phi' = c_i \log \Phi + \log k_i,$$

where k_i is an integration constant. By exponentiation, we obtain

$$\Phi' = k_i \Phi^{c_i}, \quad (3.11)$$

and here we distinguish two cases: $c_i = 1$ and $c_i \neq 1$.

When $c_i = 1$, the equation becomes $\Phi' = k_i \Phi$, which has solution $\Phi(z) = l_i \exp(k_i z)$, where l_i is another constant of integration. This corresponds to the basis function $\phi_{j_i}(x_i) = l_i k_i e^{k_i x_i}$ and it is useful to choose $l_i k_i = 1$. The corresponding basis function is then

$$\phi_{j_i}(x_i) = e^{k_i x_i}. \quad (3.12)$$

For this special choice of basis functions, $c_i = 1$ for all i and the divergence-free condition (3.3) imply that the right-hand side of (3.8) is equal to zero; hence

$$\phi_j(\mathbf{x})(t) = \phi_j(\mathbf{x})(t_0),$$

i.e. the tensor basis function is constant along the EDFVF.

Next, we consider the case when $c_i \neq 1$. Integrating (3.11) between 0 and z , we obtain

$$\Phi(z)^{1-c_i} = k_i(1-c_i)z + l_i, \quad l_i = \Phi(0)^{1-c_i},$$

where, again, k_i is a constant of integration. The constants l_i depend on the choice of the value of Φ at $z = 0$. Their choice will be discussed later. Thus, we have

$$\Phi(z) = [l_i + k_i(1-c_i)z]^{1/(1-c_i)}.$$

To recover the basis functions, we differentiate $\Phi(z)$ to obtain

$$\phi_{j_i}(z) = \Phi'(z) = k_i(l_i + k_i(1-c_i)z)^{c_i/(1-c_i)}.$$

The integration constants k_i can be chosen such that $k_i(1-c_i) = 1$. Moreover, it is useful to introduce $c_i = m_i/(1+m_i)$. With this choice, we have

$$\phi_{j_i}(x_i) = (m_i + 1)(l_i + x_i)^{m_i}, \quad m_i \neq -1.$$

In particular, if $l_i = 0$ and $m_i = j_i \in \mathbb{Z}$, we recover the monomial basis studied at great length in Xue & Zanna (2013). The case $l_i \neq 0$ is just a *shifted* monomial basis. In particular, the above formula indicates that it is also possible to use noninteger values of m_i . Note that the value $m_i = -1$ is not allowed. This value of the exponent (corresponding to $\int 1/x = \log x$) has also been discussed in Xue & Zanna (2013).

The general case $g(t, y)$, $\partial_y g \neq 0$, is not solvable for all \mathbf{a} satisfying (3.3). The argument is the following. Assume that

$$\sum_{i=1}^n a_i \frac{\phi'_{j_i}(x_i) \int \phi_{j_i}(x_i) dx_i}{\phi_{j_i}(x_i)^2} = g(t, \phi_{\mathbf{j}}(\mathbf{x})).$$

By differentiation with respect to x_i , we obtain

$$a_i \frac{\partial}{\partial x_i} \frac{\phi'_{j_i}(x_i) \int \phi_{j_i}(x_i) dx_i}{\phi_{j_i}(x_i)^2} = \partial_y g(t, \phi_{\mathbf{j}}(\mathbf{x})) \phi'_{j_i}(x_i) \prod_{l \neq i} \phi_{j_l}(x_l).$$

It is readily seen that the left-hand side depends solely on x_i , while the right-hand side also depends on all the other variables x_l , $l \neq i$. Hence, the above equation cannot have a nontrivial solution, unless for very specific choices of \mathbf{a} . \square

REMARK 3.6 In fact, the previous theorem classifies all cases of diagonalizable tensor basis functions having a closed-form solution: the monomials and the exponential basis. Other choices might still be integrable, but they might not necessarily give a closed-form solution for all \mathbf{a} satisfying (3.3). As an example, the cosine basis $\phi_{j_i}(x_i) = \cos j_i x_i$ and the sine basis $\phi_{j_i}(x_i) = \sin j_i x_i$ do not give rise to a solvable tensor basis for all \mathbf{a} , as (3.7) is not satisfied. Yet, we know that the corresponding EDFVF is integrable by Theorem 3.3 away from critical points; see also Fig. 1. Combinations of sines and cosines can be treated with solvable EDFVFs by transforming into an exponential basis (Fourier basis), as described in the next section.

REMARK 3.7 Note that the two classes of tensor basis functions satisfying Theorem 3.5, monomials and exponentials, are precisely the classes of functions that are closed under differentiation, i.e. any partial derivative of a monomial/exponential is still a monomial/exponential. Further, they are also closed under

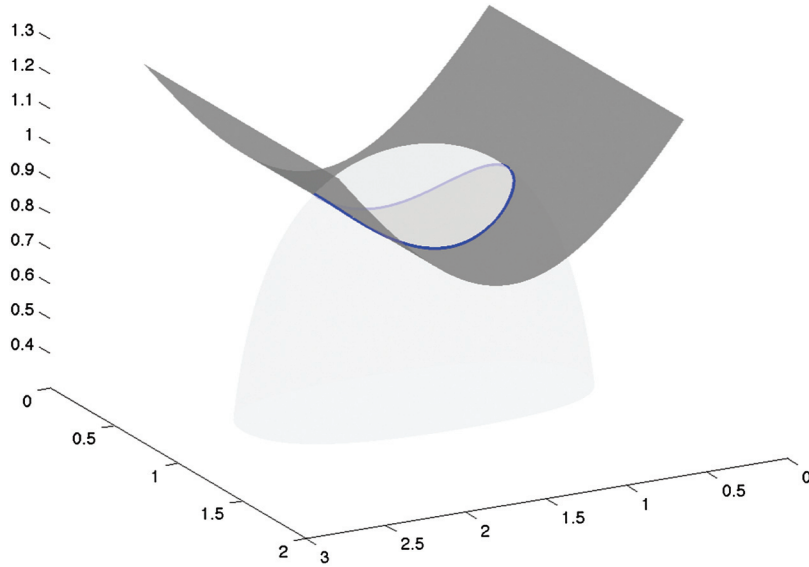


FIG. 1. A three-dimensional system with tensor basis function $\phi_{\mathbf{j}}(x_1, x_2, x_3) = \cos(j_1 x_1) \sin(j_2 x_2) \sin(j_3 x_3)$ with $\mathbf{j} = (1, 3, -1)^T$, $\mathbf{a} = (1, -2, 1)^T$ and $\mathbf{b}_2 = \mathbf{a} \times \mathbf{1}$ (cross product). The solution of the system lies on the intersection of the surfaces I_1 (lighter surface) and I_2 (darker surface).

multiplication: the product of monomials is still a monomial, the product of exponentials can still be written as a single exponential. This has the important consequence that the *commutator* of these vector fields will still be a vector field of the same type, for which the same algebraic formula holds. This is of relevance for the construction of higher-order integrators. Typically, symmetric compositions, using only the splitting of the vector field, are used. Here, we might increase the order by also using the commutators.

3.1 Connection with the splitting in shears of Quispel and McLachlan

Quispel and McLachlan introduced the technique of splitting (2.1) in divergence-free vector fields of the type

$$\dot{\mathbf{x}} = \mathbf{a}f(\mathbf{j}^T \mathbf{x}), \quad \mathbf{j}^T \mathbf{a} = 0 \quad (3.13)$$

(splitting into shears), with f an arbitrary scalar function; see McLachlan & Quispel (2004) and McLachlan *et al.* (2008). Here, \mathbf{j} need not be a vector of integers. It is easy to verify that $\mathbf{j}^T \mathbf{x} = \text{const}$ along the flow of (3.13); hence the solution is a simple translation parallel to the vector \mathbf{a} . Since $\dot{\mathbf{x}}$ is parallel to \mathbf{a} , this also implies that $\mathbf{b}^T \mathbf{x} = \text{const}$ for any \mathbf{b} orthogonal to \mathbf{a} .

PROPOSITION 3.8 The splitting into shears (3.13) is equivalent to a diagonal tensor basis splitting, with exponential basis choice (3.12).

Proof. Set $f(\mathbf{j}^T \mathbf{x}) = Ke^{\mathbf{j}^T \mathbf{x}}$ (which is always possible, since $\mathbf{j}^T \mathbf{x}$ is constant along the flow). Then, we can rewrite (3.13) as

$$\dot{x}_i = Ka j_i \frac{1}{j_i} e^{\mathbf{j}^T \mathbf{x}}, \quad i = 1, \dots, n. \quad (3.14)$$

Above, we have tacitly assumed that $j_i \neq 0$ for $i = 1, \dots, n$. This is no real restriction as, if $j_i = 0$ for some i , the remaining variables do not depend on j_i and the system is reducible.

Note that the condition $\mathbf{a}^T \mathbf{j} = 0$ is now equivalent to $\tilde{\mathbf{a}}^T \mathbf{1} = 0$, where $\tilde{\mathbf{a}} = K \mathbf{a} \otimes \mathbf{j}$ (tensor product). Thus, we see that (3.14) is precisely of the form (3.2) with basis functions $\phi_{j_i}(x_i) = e^{j_i x_i}$ and coefficient vector $\tilde{\mathbf{a}}$ obeying (3.3). Thus, Theorem 3.3 holds, and for any vector $\tilde{\mathbf{b}}$ orthogonal to $\tilde{\mathbf{a}}$, the quantity

$$\prod_{i=1}^n \left(\int \phi_{j_i}(x_i) dx_i \right)^{\tilde{b}_i}, \quad l = 1, \dots, n-1$$

is an integral. In particular, by choosing $\tilde{\mathbf{b}} = \mathbf{1}$ we recover the condition $\mathbf{j}^T \mathbf{x} = \text{const}$. For the other choices of $\tilde{\mathbf{b}}$, taking the logarithm, we see that $\sum_i \tilde{b}_i (j_i x_i - j_i)$; hence $\sum_i \tilde{b}_i j_i x_i$, is a constant of integration. Next, set $b_i = \tilde{b}_i j_i$. We have thus proved that $\frac{d}{dt} \mathbf{b}^T \mathbf{x} = 0$. Moreover, the condition $\tilde{\mathbf{b}}^T \tilde{\mathbf{a}} = 0$ is equivalent to $\mathbf{b}^T \mathbf{a} = 0$, and the condition $\tilde{\mathbf{b}}^T \mathbf{1} = 0$ is equivalent to $\mathbf{b}^T \mathbf{j} = 0$, and thus there is complete equivalence of the two formulations. \square

4. Trigonometric polynomials and Fourier bases

We next turn our attention to the case of trigonometric polynomials, which often arise when the vector field is expanded in a Fourier series. A typical application is the study of the dynamics of incompressible inviscid fluid flows. Some practical examples, often written in their real form using trigonometric polynomials, are the Arnold–Beltrami–Childress flow (ABC flow, Ricca, 2001), a prototype for the study of turbulence, and

$$\begin{aligned} \dot{x}_1 &= -\cos(\pi x_1) \sin(\pi x_2) + \varepsilon \sin(2\pi x_1) \sin(\pi x_3), \\ \dot{x}_2 &= \sin(\pi x_1) \cos(\pi x_2) + \varepsilon \sin(2\pi x_2) \sin(\pi x_3), \\ \dot{x}_3 &= 2\varepsilon \cos(\pi x_3) [\cos(2\pi x_1) + \cos(2\pi x_2)], \end{aligned}$$

a model for the study of the mixing in a laminar vortex flow with rigid boundaries, occurring in many geophysical, industrial and biophysical applications; see Solomon & Mezić (2003). For our analysis, it is more convenient to work in the complex setting. Trigonometric polynomials can always be written in a complex exponential form using the Euler formulas (4.5).

4.1 The method of expansion in a Fourier series

Consider again a vector field $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ and assume that each of the function components $f_i(\mathbf{x})$ is expanded in a Fourier series. The result will be a right-hand side that is the product of one-dimensional Fourier series in each variable, i.e.

$$\dot{x}_i = f_i(\mathbf{x}) = \sum_{\mathbf{j}} (\mathbf{c}_{\mathbf{j}})_i e^{i \mathbf{j}^T \mathbf{x}}, \quad i = 1, \dots, n, \quad (4.1)$$

where $i^2 = -1$ defines the imaginary unit. The divergence is given as

$$\nabla \cdot \mathbf{f} = i \sum_{i=1}^n \sum_{\mathbf{j}} (\mathbf{c}_{\mathbf{j}})_i j_i e^{i \mathbf{j}^T \mathbf{x}} = i \sum_{\mathbf{j}} \mathbf{c}_{\mathbf{j}}^T \mathbf{j} e^{i \mathbf{j}^T \mathbf{x}}. \quad (4.2)$$

We note at once that if the vector field \mathbf{f} is expressed in a Fourier basis, then the divergence is also expressed in a Fourier basis. To put this in the formalism of the earlier sections, we identify

$$\phi_{j_i}(x_i) = e^{ij_i x_i}, \quad \phi_{\mathbf{j}}(\mathbf{x}) = e^{i\mathbf{j}^T \mathbf{x}},$$

so that the relation between the vectors $\mathbf{a}_{\mathbf{j}}$ in Proposition 3.1 and $\mathbf{c}_{\mathbf{j}}$ is the normalization

$$(\mathbf{a}_{\mathbf{j}})_i = \frac{1}{ij_i} (\mathbf{c}_{\mathbf{j}})_i.$$

COROLLARY 4.1 Under the assumptions above, \mathbf{f} divergence-free implies $\mathbf{c}_{\mathbf{j}}^T \mathbf{j} = 0$ for any multi-index \mathbf{j} , or, equivalently, $\mathbf{a}_{\mathbf{j}}^T \mathbf{1} = 0$, in the nD Fourier expansion of \mathbf{f} .

We state the above result as a corollary because it was proved by Quispel & McLaren (2003) by direct computation on trigonometric polynomials. It is a direct consequence of Theorem 3.3.

Similarly to the approach presented for polynomials, we split the vector field into EDFVFs, each $\mathbf{F}_{\mathbf{j}}$ associated to a tensor basis element $\phi_{\mathbf{j}}$. Because of Theorem 3.5, the basis function is solvable. In fact, we have

$$\frac{d}{dt} \mathbf{j}^T \mathbf{x} = \mathbf{j}^T \dot{\mathbf{x}} = (\mathbf{j}^T \mathbf{c}_{\mathbf{j}}) e^{i\mathbf{j}^T \mathbf{x}} = 0,$$

since $\mathbf{c}_{\mathbf{j}}$ and \mathbf{j} are orthogonal (cf. the divergence-free condition). This implies that (4.1) decomposes in EDFVFs of the form $\dot{\mathbf{x}} = \mathbf{F}_{\mathbf{j}}(\mathbf{x}) = \mathbf{c}_{\mathbf{j}} e^{i\mathbf{j}^T \mathbf{x}}$, whose solution is given explicitly as

$$\mathbf{x}(t_0 + h) = \mathbf{x}_0 + h \mathbf{c}_{\mathbf{j}} e^{i\mathbf{j}^T \mathbf{x}_0}. \quad (4.3)$$

A naive implementation of this method would require complex arithmetic. However, if the function \mathbf{f} is real, then the coefficients of the Fourier transform obey the condition

$$\mathbf{c}_{-\mathbf{j}} = \bar{\mathbf{c}}_{\mathbf{j}}$$

(the bar denotes complex conjugation). This implies that the EDFVFs $\mathbf{F}_{\mathbf{j}}$ and $\mathbf{F}_{-\mathbf{j}}$ share the same integrals and can be combined together. Thus, we need to sum over just half of the multi-indices and, instead of (4.1), we can consider

$$\dot{\mathbf{x}} = \tilde{F}_{\mathbf{j}}(\mathbf{x}) = \mathbf{c}_{\mathbf{j}} e^{i\mathbf{j}^T \mathbf{x}} + \bar{\mathbf{c}}_{\mathbf{j}} e^{-i\mathbf{j}^T \mathbf{x}}. \quad (4.4)$$

Writing $\mathbf{c}_{\mathbf{j}} = \alpha_{\mathbf{j}} + i\beta_{\mathbf{j}}$, and using the identities

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i}, \quad \cos z = \frac{e^{iz} + e^{-iz}}{2}, \quad (4.5)$$

it is readily seen that (4.4) becomes

$$\dot{\mathbf{x}} = 2\alpha_{\mathbf{j}} \cos(\mathbf{j}^T \mathbf{x}) - 2\beta_{\mathbf{j}} \sin(\mathbf{j}^T \mathbf{x}), \quad (4.6)$$

which can be integrated exactly as

$$\mathbf{x}(t + h) = \mathbf{x}_0 + h(2\alpha_{\mathbf{j}} \cos(\mathbf{j}^T \mathbf{x}_0) - 2\beta_{\mathbf{j}} \sin(\mathbf{j}^T \mathbf{x}_0)). \quad (4.7)$$

We recognize (4.7) as the method proposed in Quispel & McLaren (2003).

5. Tensor product of Fourier series and the monomial basis

So far, the pure monomial choice reduces to the methods in [Xue & Zanna \(2013\)](#) and the pure trigonometric choice reduces to the methods of [Quispel & McLaren \(2003\)](#). However, because of the generality of Proposition 2.1 we are also able to treat the mixed polynomial–trigonometric case, which one could not deal with systematically before.

We start from the divergence function, which we assume can be written in the form

$$\nabla \cdot \mathbf{f} = \sum_{\mathbf{j}} p_{\mathbf{j}} \mathbf{x}^{\mathbf{j}_m} e^{i \mathbf{j}_f^T \mathbf{x}},$$

where, by an appropriate relabelling of the variables, we can assume

$$\mathbf{j} = \mathbf{j}_m \oplus \mathbf{j}_f, \quad \mathbf{j}_m = \begin{pmatrix} j_1 \\ \vdots \\ j_p \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{j}_f = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ j_{p+1} \\ \vdots \\ j_n \end{pmatrix}.$$

Here, the subscript m is associated to the projection onto the first p components (monomial part), while the subscript f refers to the projection onto the remaining $n - p$ components (Fourier part). Dropping the dependence of the coefficient \mathbf{a} on the index \mathbf{j} , the corresponding EDFVF is

$$\dot{\mathbf{x}} = \mathbf{F}_{\mathbf{j}}(\mathbf{x}) = \mathbf{a} \otimes \begin{pmatrix} \frac{x_1}{j_1 + 1} \\ \vdots \\ \frac{x_p}{j_p + 1} \\ \mathbf{i} \\ \frac{1}{j_{p+1}} \\ \vdots \\ \mathbf{i} \\ \frac{1}{j_n} \end{pmatrix} \mathbf{x}^{\mathbf{j}_m} e^{i \mathbf{j}_f^T \mathbf{x}}, \quad \mathbf{a}^T \mathbf{1} = 0 \quad (5.1)$$

(the \otimes denoting the tensor (componentwise) product of two vectors). Equivalently, we could write the EDFVF using the unscaled coefficients \mathbf{c} ,

$$\dot{\mathbf{x}} = \mathbf{F}_{\mathbf{j}}(\mathbf{x}) = \mathbf{c} \otimes \begin{pmatrix} x_1 \\ \vdots \\ x_p \\ 1 \\ \vdots \\ 1 \end{pmatrix} \mathbf{x}^{\mathbf{j}_m} e^{i \mathbf{j}_f^T \mathbf{x}}, \quad (5.2)$$

subject to

$$\mathbf{c}^T((\mathbf{j} + \mathbf{1})_m + \mathbf{i}\mathbf{j}_f) = 0. \quad (5.3)$$

From Theorem 3.5, it follows that the basis function $\phi_{\mathbf{j}}(\mathbf{x}) = \mathbf{x}^{\mathbf{j}_m} e^{\mathbf{i}\mathbf{j}_f^T \mathbf{x}}$ obeys the differential equation

$$\frac{d}{dt} \phi_{\mathbf{j}}(\mathbf{x}(t)) = C \phi_{\mathbf{j}}(\mathbf{x})^2, \quad C = \mathbf{c}^T(\mathbf{j}_m + \mathbf{i}\mathbf{1}_f),$$

with solution

$$\phi_{\mathbf{j}}(\mathbf{x}) = \phi_{\mathbf{j}}(\mathbf{x}_0)(1 - C(t - t_0)\phi_{\mathbf{j}}(\mathbf{x}_0))^{-1}.$$

Now, for $i = 1, \dots, p$, we have

$$\dot{x}_i = c_i x_i \phi_{\mathbf{j}}(t),$$

with solution

$$x_i(t) = x_i(t_0)(1 - C\phi_{\mathbf{j}}(\mathbf{x}_0)(t - t_0))^{-r_i}, \quad r_i = \frac{c_i}{C}, \quad i = 1, \dots, p, \quad (5.4)$$

while, for $i = p + 1, \dots, n$ we have

$$\dot{x}_i = c_i \phi_{\mathbf{j}}(t),$$

with solution

$$x_i(t) = x_i(t_0) - r_i \log(1 - C\phi_{\mathbf{j}}(\mathbf{x}_0)(t - t_0)), \quad i = p + 1, \dots, n, \quad (5.5)$$

and r_i as in (5.4).

5.1 Avoiding complex arithmetic

In the case when basis elements $\mathbf{x}^{\mathbf{j}_m} e^{\mathbf{i}\mathbf{j}_f^T \mathbf{x}}$ are used, the coefficients are generally complex and the method, as described before, will require complex arithmetic. This might be a problem, because the logarithm functions in (5.5) will require the tracking of the argument function, for a correct complex part. As most applications involve real coefficients, it is a disadvantage (and computationally more expensive) to use complex arithmetic.

A first attempt would be to try to generalize the techniques for pure trigonometric tensor products described earlier (and in Quispel & McLaren, 2003). Let $\mathbf{c} = \boldsymbol{\alpha} + \mathbf{i}\boldsymbol{\beta}$. It is readily seen that the divergence-free condition (5.3) is equivalent to the two conditions

$$\boldsymbol{\alpha}^T(\mathbf{j} + \mathbf{1})_m - \boldsymbol{\beta}^T \mathbf{j}_f = 0, \quad (5.6)$$

$$\boldsymbol{\beta}^T(\mathbf{j} + \mathbf{1})_m + \boldsymbol{\alpha}^T \mathbf{j}_f = 0. \quad (5.7)$$

These are, in fact, the conditions for the vector fields

$$T_1 : \begin{cases} \dot{x}_i = \alpha_i x_i \mathbf{x}^{\mathbf{j}_m} \cos(\mathbf{j}_f^T \mathbf{x}), \\ \dot{x}_i = -\beta_i \mathbf{x}^{\mathbf{j}_m} \sin(\mathbf{j}_f^T \mathbf{x}), \end{cases} \quad T_2 : \begin{cases} \dot{x}_i = -\beta_i x_i \mathbf{x}^{\mathbf{j}_m} \sin(\mathbf{j}_f^T \mathbf{x}), \\ \dot{x}_i = \alpha_i \mathbf{x}^{\mathbf{j}_m} \cos(\mathbf{j}_f^T \mathbf{x}), \end{cases} \quad i = 1, \dots, p, \quad i = p + 1, \dots, n, \quad (5.8)$$

to be volume preserving. The vector field $T_1 + T_2$ can be recognized as the real part of (5.2) (we assume that the solution is real). Note the resemblance to (4.6), which we treated simultaneously (because the sine and cosine part shared the same integral). Here, the vector field T_1 has integral $\mathbf{x}^{(\mathbf{j}+\mathbf{1})_m} \sin(\mathbf{j}_f^T \mathbf{x})$, while the vector field T_2 has integral $\mathbf{x}^{(\mathbf{j}+\mathbf{1})_m} \cos(\mathbf{j}_f^T \mathbf{x})$. These two integrals do not commute, indicating that there is no easy way to solve $T_1 + T_2$ simultaneously. On the other hand, we observe that

$\mathbf{x}^{\mathbf{j}_m} \cos(\mathbf{j}_f^T \mathbf{x})$ and $\mathbf{x}^{\mathbf{j}_m} \sin(\mathbf{j}_f^T \mathbf{x})$ are also basis functions (though not tensor product, but linear combinations of tensor bases). Thus, a natural strategy is to attempt solving T_1 and T_2 separately.

We set

$$\begin{aligned} T_1 : \phi_m(\mathbf{x}) &= \mathbf{x}^{\mathbf{j}_m}, & \phi_f(\mathbf{x}) &= \cos(\mathbf{j}_f^T \mathbf{x}), \\ T_2 : \phi_m(\mathbf{x}) &= \mathbf{x}^{\mathbf{j}_m}, & \phi_f(\mathbf{x}) &= \sin(\mathbf{j}_f^T \mathbf{x}). \end{aligned}$$

It is readily observed that

$$\begin{aligned} \phi'_m &= \kappa_1 \phi_m^2 \phi_f, \\ \phi'_f &= \kappa_2 \phi_m (1 - \phi_f^2), \end{aligned} \tag{5.9}$$

with

$$\begin{aligned} T_1 : \kappa_1 &= \boldsymbol{\alpha}^T \mathbf{j}_m, & \kappa_2 &= \boldsymbol{\beta}^T \mathbf{j}_f, \\ T_2 : \kappa_1 &= -\boldsymbol{\beta}^T \mathbf{j}_m, & \kappa_2 &= \boldsymbol{\alpha}^T \mathbf{j}_f. \end{aligned} \tag{5.10}$$

The two equations in (5.9) can be combined into

$$\frac{\phi'_m}{\phi_m} = \frac{\kappa_1}{\kappa_2} \frac{\phi'_f \phi_f}{1 - \phi_f^2},$$

which can be integrated by separation of variables to give

$$\frac{\phi_m(t)}{\phi_m(t_0)} = \left(\frac{1 - \phi_f^2(t)}{1 - \phi_f^2(t_0)} \right)^{-\kappa_1/2\kappa_2}.$$

Setting

$$\rho = \frac{\kappa_1}{2\kappa_2}, \quad K = \kappa_2 \phi_m(t_0) (1 - \phi_f^2(t_0))^\rho$$

and substituting $\phi_m(t)$ into ϕ'_f , we obtain the differential equation

$$\frac{\phi'_f}{(1 - \phi_f^2)^{1-\rho}} = K, \tag{5.11}$$

which can be solved by separation of variables. The solution is given in terms of hypergeometric functions,

$$\phi_f {}_2F_1 \left[\begin{matrix} \frac{1}{2}, 1 - \rho; \\ \frac{3}{2}; \end{matrix} \phi_f^2 \right] \bigg|_{\phi_f(t_0)}^{\phi_f(t)} = K(t - t_0); \tag{5.12}$$

see [Abramowitz & Stegun \(1965\)](#) and [Rainville \(1967\)](#). Some special cases of the parameter ρ are worth discussion.

Case $\rho = 0$ ($\kappa_1 = 0, \kappa_2 \neq 0$). In this case, directly from (5.9), we have $\phi'_m = 0$. This gives

$$\begin{aligned}\phi_m(t) &= \phi_m(t_0), \\ \phi_f(t) &= -\frac{1 - K_1 e^{2K(t-t_0)}}{1 + K_1 e^{2K(t-t_0)}}, \quad K_1 = \frac{1 + \phi_f(t_0)}{1 - \phi_f(t_0)}, \quad K = \kappa_2 \phi_m(t_0).\end{aligned}$$

Next, we integrate the monomial part in (5.8) to obtain

$$T_1 : x_i(t) = x_i(t_0) e^{-\alpha_i \phi_m(t_0)(t-t_0)} \left[\frac{1 + K_1 e^{2K(t-t_0)}}{1 + K_1} \right]^{\alpha_i \phi_m(t_0)K}, \quad (5.13)$$

$$T_2 : x_i(t) = x_i(t_0) e^{\beta_i \phi_m(t_0)(t-t_0)} \left[\frac{1 + K_1 e^{2K(t-t_0)}}{1 + K_1} \right]^{-\beta_i \phi_m(t_0)K}, \quad i = 1, \dots, p, \quad (5.14)$$

and, for the Fourier part,

$$T_1 : x_i(t) = x_i(t_0) - \beta_i s_0 \frac{(1 + K_1)}{\kappa_2 \sqrt{K_1}} \tan^{-1} \left[\frac{K_1 + \sqrt{K_1} e^{K(t-t_0)}}{1 - K_1 \sqrt{K_1}} \right], \quad (5.15)$$

$$T_2 : x_i(t) = x_i(t_0) + \alpha_i c_0 \frac{(1 + K_1)}{\kappa_2 \sqrt{K_1}} \tan^{-1} \left[\frac{K_1 + \sqrt{K_1} e^{K(t-t_0)}}{1 - K_1 \sqrt{K_1}} \right], \quad i = p + 1, \dots, n, \quad (5.16)$$

where

$$s_0 = \sin(\mathbf{j}_f^T \mathbf{x}(t_0)), \quad c_0 = \cos(\mathbf{j}_f^T \mathbf{x}(t_0)).$$

Case $\rho = \infty$ ($\kappa_1 \neq 0, \kappa_2 = 0$). Directly from (5.9), we have $\phi'_f = 0$. In particular, this implies that both $\cos(\mathbf{j}_f^T \mathbf{x}) = c_0$ and $\sin(\mathbf{j}_f^T \mathbf{x}) = s_0$. The monomial part is closed and can be solved by the methods described in Xue & Zanna (2013) (see also (5.4)):

$$T_1 : x_i(t) = x_i(t_0) (1 - \kappa_1 c_0 \phi_m(t_0)(t - t_0))^{-r_i}, \quad r_i = \frac{\alpha_i}{\kappa_1},$$

$$T_2 : x_i(t) = x_i(t_0) (1 - \kappa_1 s_0 \phi_m(t_0)(t - t_0))^{-r_i}, \quad r_i = \frac{-\beta_i}{\kappa_1}, \quad i = 1, \dots, p$$

and

$$T_1 : x_i(t) = x_i(t_0) - q_i \log(1 - \kappa_1 c_0 \phi_m(t_0)(t - t_0)), \quad q_i = \frac{-\beta_i s_0}{\kappa_1 c_0},$$

$$T_2 : x_i(t) = x_i(t_0) - q_i \log(1 - \kappa_1 s_0 \phi_m(t_0)(t - t_0)), \quad q_i = \frac{\alpha_i c_0}{\kappa_1 s_0}, \quad i = p + 1, \dots, n$$

for the Fourier part.

Case $\rho = \frac{0}{0}$ ($\kappa_1 = \kappa_2 = 0$). In this case, $\phi'_m = \phi'_f = 0$. As above, we have $\cos(\mathbf{j}_f^T \mathbf{x}) = c_0$ and $\sin(\mathbf{j}_f^T \mathbf{x}) = s_0$. The solutions are easily found:

$$T_1 : x_i(t) = x_i(t_0) e^{\alpha_i c_0 \phi_m(t_0)(t-t_0)},$$

$$T_2 : x_i(t) = x_i(t_0) e^{-\beta_i s_0 \phi_m(t_0)(t-t_0)}, \quad i = 1, \dots, p$$

for the monomial part, and

$$\begin{aligned} T_1 : x_i(t) &= x_i(t_0) - \beta_i s_0 \phi_m(t_0)(t - t_0), \\ T_2 : x_i(t) &= x_i(t_0) + \alpha_i c_0 \phi_m(t_0)(t - t_0), \quad i = p + 1, \dots, n \end{aligned}$$

for the Fourier part.

Case $\rho = 1$ ($\kappa_1 = 2\kappa_2 \neq 0$). In this case, (5.11) reduces to $\phi_f' = K$. This gives

$$\begin{aligned} \phi_f(t) &= \phi_f(t_0) + K(t - t_0), \\ \phi_m(t) &= \phi_m(t_0) \left[\frac{1 - \phi_f(t)^2}{1 - \phi_f^2(t_0)} \right]^{-1}. \end{aligned}$$

Thus, we have

$$x_i(t) = x_i(t_0) \left[\frac{1 - \phi_f(t)^2}{1 - \phi_f^2(t_0)} \right]^{-r_i}, \quad T_1 : r_i = \frac{\alpha_i}{2\kappa_2}, \quad T_2 : r_i = \frac{-\beta_i}{2\kappa_2}, \quad i = 1, \dots, p$$

for the monomial part. For the Fourier part, we focus on T_1 and observe that $y = \sin(\mathbf{j}_f^T \mathbf{x})$ obeys the same differential equation as x_i of the monomial part upon replacement of α_i by $-\kappa_2$. Thus, it has a similar solution:

$$\sin(\mathbf{j}_f^T \mathbf{x}) = s_0 \left[\frac{1 - \phi_f(t)^2}{1 - \phi_f^2(t_0)} \right]^{1/2}.$$

A similar formula holds for T_2 . Substituting into \dot{x}_i and integrating, we have

$$\begin{aligned} x_i &= x_i(t_0) + \frac{q_i \phi_m(t_0)}{K} (\sin^{-1} \phi_f - \sin^{-1} \phi_f(t_0)) \\ &= x_i(t_0) - \frac{q_i \phi_m(t_0)}{K} (\cos^{-1} \phi_f - \cos^{-1} \phi_f(t_0)), \end{aligned}$$

where

$$T_1 : q_i = -\beta_i s_0 |s_0|, \quad T_2 : q_i = \alpha_i c_0 |c_0|.$$

The general case requires the inversion of the special function in (5.12). Note that if $\rho > 1$, the series terminates (finite sum).

6. Divergence-free with measure $\mu(\mathbf{x}) \neq 1$

In the above, we have discussed the case of how to split divergence-free methods for the case of *canonical* variables, i.e. if the vector field \mathbf{f} obeys

$$\nabla \cdot \mathbf{f} = 0.$$

In some cases, the divergence-free condition is given with respect to a measure $\mu(\mathbf{x})$, meaning that the vector field $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ obeys the condition

$$\nabla \cdot (\mu \mathbf{f}) = 0. \quad (6.1)$$

The scope of this section is to extend the results for the canonical case $\mu = 1$ to the generic case when $\mu = \mu(\mathbf{x})$. A result very similar to the case $\mu = 1$ is true also in this case.

PROPOSITION 6.1 Let $\phi_{\mathbf{j}}(\mathbf{x})$ be in the form (3.1) for some arbitrary tensor-product basis. Consider a decomposition

$$\nabla \cdot (\mu \mathbf{f}) = \sum_{\mathbf{j}} p_{\mathbf{j}} \phi_{\mathbf{j}}(\mathbf{x}).$$

The EDFVF $\mathbf{F}_{\mathbf{j}}$ (2.5) corresponding to the multi-index $\mathbf{j} = (j_1, \dots, j_n)$ must be of the form

$$\dot{x}_i = (a_{\mathbf{j}})_i \frac{1}{\mu(\mathbf{x})} \int \phi_{j_i}(x_i) dx_i \prod_{\substack{l=1 \\ l \neq i}}^n \phi_{j_l}(x_l) \quad (6.2)$$

$$= (a_{\mathbf{j}})_i \frac{\int \phi_{j_i}(x_i) dx_i}{\phi_{j_i}(x_i)} \frac{\phi_{\mathbf{j}}(\mathbf{x})}{\mu(\mathbf{x})}, \quad i = 1, \dots, n, \quad (6.3)$$

where $\mathbf{a}_{\mathbf{j}}^T \mathbf{1} = 0$, as in (3.3). Furthermore, for any \mathbf{b} orthogonal to $\mathbf{a}_{\mathbf{j}}$ and $\mathbf{1}$, one has that

$$\left(\int \phi_{j_1}(x_1) dx_1 \right)^{b_1} \cdots \left(\int \phi_{j_n}(x_n) dx_n \right)^{b_n}$$

is a conserved quantity of the system.

The proof is analogous to the case $\mu = 1$. Thus, also, the generic case is integrable, provided that the integrals are linearly independent.

PROPOSITION 6.2 The vector field (6.3) is equivalent to the vector field (3.2) upon time reparametrization $\frac{dt}{d\tau} = \mu(\mathbf{x})$.

Thus, all the knowledge of the canonical case $\mu = 1$ can in principle be extended to the more general case $\mu \neq 1$.

7. Concluding remarks

We have presented a framework for generating volume-preserving splitting methods based on a decomposition of the divergence of the vector field using an appropriate tensor product of basis functions of one variable. The vector field is then decomposed into the sum of EDFVFs, each of them corresponding to a basis function. This theory includes, as a special case, the monomial basis of Xue & Zanna (2013) and the trigonometric splitting of Quispel & McLaren (2003). We have introduced the concept of diagonalizable EDFVFs and classified solvable ones: the monomial basis and the exponential basis. We have further shown that the method of splitting into shears proposed by McLachlan & Quispel (2004) is equivalent to using a special class of diagonalizable EDFVFs, corresponding to the choice of exponential basis.

In addition to giving a unifying view of some types of volume-preserving splitting methods already known in the literature, the present approach allows us to give a closed-form solution also to other types of vector fields that could not be treated before, namely those corresponding to the mixed tensor product of monomial and exponential (including trigonometric) basis functions.

Funding

This work has been supported by the Norwegian Research Council through the project ‘Geometric Numerical Integration and Applications’, NFR project no. 191178/V30. Funding to pay the Open Access publication charges for this article was provided by the University of Bergen.

REFERENCES

- ABRAMOWITZ, M. & STEGUN, I. A. (1965) *Handbook of Mathematical Functions*. New York: Dover.
- ARNOLD, V. I. & KHESIN, B. A. (1998) *Topological Methods in Hydrodynamics*. New York: Springer.
- CHARTIER, P. & MURUA, A. (2007) Preserving first integrals and volume forms of additively split systems. *IMA J. Numer. Anal.*, **27**, 381–405.
- ISERLES, A., QUISPTEL, G. R. W. & TSE, P. S. P. (2007) B-series methods cannot be volume-preserving. *BIT*, **47**, 351–378.
- KANG, F. & SHANG, Z. J. (1995) Volume-preserving algorithms for source-free dynamical systems. *Numer. Math.*, **71**, 451–463.
- MCLACHLAN, R. I., MUNTHE-KAAS, H. Z., QUISPTEL, G. R. W. & ZANNA, A. (2008) Explicit volume-preserving splitting methods for linear and quadratic divergence-free vector fields. *Found. Comput. Math.*, **8**, 335–355.
- MCLACHLAN, R. I. & SCOVEL, C. (1998) A survey of open problems in symplectic integration. *Fields Inst. Commun.*, **10**, 151–180.
- MCLACHLAN, R. & QUISPTEL, G. (2004) Explicit geometric integration of polynomial vector fields. *BIT*, **44**, 515–538.
- NEAL, R. M. (2011) MCMC using Hamiltonian dynamics. *Handbook of Markov Chain Monte Carlo* (S. Brooks, A. Gelman, G. L. Jones & X.-L. Meng eds). Boca Raton, FL: Chapman & Hall/CRC.
- QIN, M.-Z. & ZHU, W.-J. (1993) Volume-preserving schemes and numerical experiments. *Comput. Math. Appl.*, **26**, 33–42.
- QUISPTEL, G. R. W. (1995) Volume-preserving integrators. *Phys. Lett. A*, **206**, 26–30.
- QUISPTEL, G. & McLAREN, D. (2003) Explicit volume-preserving and symplectic integrators for trigonometric polynomial flows. *J. Comp. Phys.*, **186**, 308–316.
- RAINVILLE, D. (1967) *Special Functions*. New York: Macmillan.
- RICCA, R. (2001) *An Introduction to the Geometry and Topology of Fluid Flows*. Dordrecht, Boston: Kluwer Academic Publishers.
- SCOVEL, J. C. (1991) Symplectic numerical integration of Hamiltonian systems. *The Geometry of Hamiltonian Systems* (T. Ratiu ed.), MSRI Series, vol. 22. New York: Springer, pp. 463–496.
- SHANG, Z. J. (1994) Construction of volume-preserving difference schemes for source-free systems via generating functions. *J. Comput. Math.*, **12**, 265–272.
- SOLOMON, T. H. & MEZIĆ, I. (2003) Uniform resonant chaotic mixing in fluid flows. *Nature*, **524**, 376–380.
- XUE, H. & ZANNA, A. (2013) Explicit volume-preserving splitting methods for polynomial divergence-free vector fields. *BIT Numer. Math.*, **53**, 265–281.
- XUE, H. & ZANNA, A. (2014) Generating functions and volume preserving mappings. *Discrete Contin. Dyn. Syst. Ser. A*, **34**, 1229–1249.