Manuscript submitted to
AIMS' Journals
Volume **X**, Number **0X**, XX **200X**

Website: http://AIMsciences.org

pp. **X–XX**

DIRICHLET SERIES FOR DYNAMICAL SYSTEMS OF FIRST-ORDER ORDINARY DIFFERENTIAL EQUATIONS

Bin Wang

Department of Mathematics, Nanjing University
State Key Laboratory for Novel Software Technology at Nanjing University
Nanjing 210093, P.R.China
and
Department of Applied Mathematics and Theoretical Physics
Centre for Mathematical Sciences, University of Cambridge
Wilberforce Rd, Cambridge CB3 0WA, United Kingdom

Arieh Iserles

Department of Applied Mathematics and Theoretical Physics
Centre for Mathematical Sciences, University of Cambridge
Wilberforce Rd, Cambridge CB3 0WA, United Kingdom

(Communicated by the associate editor name)

Abstract. In this paper, inspired by the work by A. Iserles and G. Söderlind [*Global bounds on numerical error for ordinary differential equations, J. Complexity, 9 (1993), pp. 97-112*], we present comprehensive discussion on Dirichlet series for dynamical systems of first-order ordinary differential equations (ODEs). We first derive the scheme of Dirichlet approximation for scalar dynamical systems and present the bounds on the terms of Dirichlet series. The global error and the right choice of a term in Dirichlet series are analysed and two numerical experiments are carried out to demonstrate the efficiency of Dirichlet approximation. Then we consider applying Dirichlet series to multivariate dynamical systems and present a new scheme of Dirichlet approximation for such systems. Some discussion and a numerical experiment are accordingly carried out for the new Dirichlet approximation. Compared with routine time-stepping algorithms, Dirichlet series does not need time stepping and yields a continuous solution that is equally valid along an interval, which is significant for obtaining long-time numerical solution. As a result of the special nature of Dirichlet series, the Dirichlet approximation delivers considerable information on dynamical systems of first-order ODEs and provides a novel and effective approach to numerical solutions of these dynamical systems.

1. **Introduction.** Dynamical systems of first-order ordinary differential equations (ODEs) are an important and basic aspect of dynamical systems ([16, 13, 14]), which arise in various fields of science and technology, such as applied mathematics, mechanics, physics, astronomy, molecular biology and engineering. Many efficient

2000 *Mathematics Subject Classification.* Primary: 37M10, 37N30; Secondary: 65L05, 65L70.

Key words and phrases. Dirichlet series, dynamical systems, first-order ordinary differential equations, stationary points, global bounds.

The first author is supported in part by the Natural Science Foundation of China under Grant 11271186, by the Specialized Research Foundation for the Doctoral Program of Higher Education under Grant 20100091110033 and by the University Postgraduate Research and Innovation Project of Jiangsu Province 2012 under Grant CXZZ12_0028.

numerical methods have been proposed for this kind of dynamical systems such as Runge–Kutta methods, extrapolation methods, multi-step methods, exponential fitting and trigonometric fitting methods and so on and we refer to [15, 3, 5, 8, 9, 11, 4, 1, 7, 2] for examples. These well-known methods are all routine time-stepping algorithms and require step by step computations to obtain an approximation to the system in a long interval. However, in this paper we devote our attention to a very different approach, which dose not need time stepping and applies a new means to solving dynamical systems of first-order ODEs. For the following scalar dynamical system of first-order ODE

$$y' = f(y), \tag{1}$$

authors in [10] considered applying *Dirichlet series* (see [6, 12]) to system (1) and they presented some results on this interesting approach. However, in [10] the method of Dirichlet series is introduced only for the scalar equation (1) with a polynomial $f(y)$ and the corresponding results are given without any proof. Authors in [10] stated that those results are both incomplete and speculative. This is not very effective for the application of Dirichlet series as an effective approach to approximating solutions of dynamical systems of first-order ODEs. Inspired by [10] and in order to improve upon it, we present in this paper a complete discussion on Dirichlet series for dynamical systems of first-order ODEs. The method of Dirichlet series preserves some properties of dynamical systems and provides numerical solution which does not need time stepping and is continuous in a long interval. This makes Dirichlet series differ form routine-time stepping algorithms.

The outline of this paper is as follows. In Section 2, we introduce the scheme of Dirichlet series for scalar dynamical systems and then seek the expansion of the solution of (1) in Dirichlet series. First few terms of Dirichlet series are derived in Section 3 and then the bounds on the terms of Dirichlet series are given in Section 4. In Section 5, the global error for Dirichlet series is estimated and the convergence of the Dirichlet expansion is discussed. Section 6 considers the right choice of σ_0 (a free term of Dirichlet series) and two numerical experiments are performed to demonstrate the efficiency of the Dirichlet approximation and support our theoretical analysis. In Section 7, we discuss the applications of Dirichlet series to multivariate dynamical systems and extend Dirichlet series to a particular multivariate dynamical system. Some results are presented and a multivariate experiment is carried out to demonstrate the efficiency of the Dirichlet approximation for multivariate dynamical systems. Section 8 presents the conclusions of this paper.

2. Dirichlet series for scalar dynamical systems.

In this section we consider the following scalar dynamical system of first-order ODE

$$y' = f(y), \qquad y(t_0) = y_0, \tag{2}$$

where $y : \mathbb{R} \to \mathbb{C}$ is the solution of (2) and $f : \mathbb{C} \to \mathbb{C}$ is an analytic function. We seek an expansion of the solution of (2) in *Dirichlet series* (see [6, 12])

$$y(t) = \sum_{n=0}^{\infty} \sigma_n e^{n\lambda t}, \qquad t \in \mathbb{R}, \tag{3}$$

where σ_1 is an arbitrary (for the time being) complex parameter, σ_0, λ ($\lambda \neq 0$) and $\{\sigma_n\}_{n=2}^{\infty}$ are independent of t and are determined by substitution of (3) into (2).

Inserting (3) into (2) yields

$$\sum_{n=0}^{\infty} n\lambda\sigma_n e^{n\lambda t} = f(y) = \sum_{l=0}^{\infty} \frac{f^{(l)}(\sigma_0)}{l!}\Big(\sum_{n=1}^{\infty}\sigma_n e^{n\lambda t}\Big)^l, \tag{4}$$

where $f^{(l)}$ denotes the lth-derivative of $f(y)$ with respect to y.

Equating the constant terms in both sides of (4) gives

$$0 = f(\sigma_0), \tag{5}$$

which shows that σ_0 must be a zero of f. In other words, σ_0 is a stationary point of the dynamical system (2). This is important because often many properties of a stationary point have been studied and we can use those results. In that case the expansion (3) is more natural and has good properties. We will discuss this point further in Section 6.

Considering the coefficients of $e^{n\lambda t}$ with $n = 1$ in both sides of (4), we get

$$\lambda = f'(\sigma_0). \tag{6}$$

For $n \geq 2$, comparing the coefficients of $e^{n\lambda t}$ in both sides of (4) yields

$$\sigma_n = \frac{1}{n\lambda}\Big[\sum_{l=1}^{\infty}\frac{f^{(l)}(\sigma_0)}{l!}\sum_{\mathbf{j}\in\mathbb{J}_{n,l}}\binom{l}{j_1, j_2, \ldots, j_n}\sigma_1^{j_1}\sigma_2^{j_2}\cdots\sigma_n^{j_n}\Big], \qquad n \geq 2, \tag{7}$$

where $J_{n,l}$ is defined as

$$\mathbb{J}_{n,l} = \{(j_1, \ldots, j_n) \in {\mathbb{Z}_+}^n : \ j_1 + j_2 + \ldots + j_n = l, \ j_1 + 2j_2 + \ldots + nj_n = n\} \tag{8}$$

and

$$\binom{l}{j_1, j_2, \ldots, j_n} = \frac{l!}{j_1!j_2!\cdots j_n!}.$$

Here $\mathbb{Z}_+$ is the set of nonnegative integers. From the formula (8), it follows immediately that $l \leq n$. Thus (7) becomes

$$\sigma_n = \frac{1}{n\lambda}\Big[\sum_{l=1}^{n}\frac{f^{(l)}(\sigma_0)}{l!}\sum_{\mathbf{j}\in\mathbb{J}_{n,l}}\binom{l}{j_1, j_2, \ldots, j_n}\sigma_1^{j_1}\sigma_2^{j_2}\cdots\sigma_n^{j_n}\Big]. \tag{9}$$

By the definition $\mathbb{J}_{n,l}$ (8), we know that $\mathbb{J}_{n,1} = \{(0, 0, \ldots, 0, 1)\}$ and $\mathbb{J}_{n,l} \equiv \widetilde{\mathbb{J}}_{n,l}$ for $l \geq 2$. Here $\widetilde{\mathbb{J}}_{n,l}$ is defined as

$$\widetilde{\mathbb{J}}_{n,l} = \{(j_1, \ldots, j_{n-1}, 0) \in \mathbb{Z}_+^n : \ j_1+j_2+\ldots+j_{n-1} = l, \ j_1+2j_2+\ldots+(n-1)j_{n-1} = n\}. \tag{10}$$

Keeping this in mind, we arrive at from (9)

$$\sigma_n = \frac{1}{n}\sigma_n + \frac{1}{n\lambda}\Big[\sum_{l=2}^{n}\frac{f^{(l)}(\sigma_0)}{l!}\sum_{\mathbf{j}\in\widetilde{\mathbb{J}}_{n,l}}\binom{l}{j_1, j_2, \ldots, j_{n-1}, 0}\sigma_1^{j_1}\sigma_2^{j_2}\cdots\sigma_{n-1}^{j_{n-1}}\Big].$$

Thus

$$\sigma_n = \frac{1}{(n-1)\lambda}\Big[\sum_{l=2}^{n}\frac{f^{(l)}(\sigma_0)}{l!}\sum_{\mathbf{j}\in\widetilde{\mathbb{J}}_{n,l}}\binom{l}{j_1, j_2, \ldots, j_{n-1}}\sigma_1^{j_1}\sigma_2^{j_2}\cdots\sigma_{n-1}^{j_{n-1}}\Big], \quad n = 2, 3, \ldots. \tag{11}$$

This is the pattern by which each σ_n for $n \geq 2$ can be obtained explicitly through a recursive procedure.

3. **The terms of Dirichlet series.** Based on (11), we derive explicitly the first few terms of Dirichlet series (3) in this section.

When $n=2$ we obtain $\widetilde{\mathbb{J}}_{2,2}=\{(2,0)\}$ and then

$$\sigma_2=\frac{1}{2\lambda}f''(\sigma_0)\sigma_1^2. \tag{12}$$

For $n=3$ we have $\widetilde{\mathbb{J}}_{3,2}=\{(1,1,0)\}$ and $\widetilde{\mathbb{J}}_{3,3}=\{(3,0,0)\}$. By (11),

$$\sigma_3=\frac{1}{2\lambda}\Big[\frac{f''(\sigma_0)}{2!}2\sigma_1\sigma_2+\frac{f^{(3)}(\sigma_0)}{3!}\sigma_1^3\Big].$$

From (12) this implies

$$\sigma_3=\frac{\sigma_1^3}{12\lambda^2}\Big[3(f''(\sigma_0))^2+f'(\sigma_0)f^{(3)}(\sigma_0)\Big]. \tag{13}$$

The case of $n=4$ yields $\widetilde{\mathbb{J}}_{4,2}=\{(1,0,1,0),(0,2,0,0)\}$, $\widetilde{\mathbb{J}}_{4,3}=\{(2,1,0,0)\}$ and $\widetilde{\mathbb{J}}_{4,4}=\{(4,0,0,0)\}$. Thus

$$\sigma_4=\frac{1}{3\lambda}\Big[\frac{f''(\sigma_0)}{2!}(2\sigma_1\sigma_3+\sigma_2^2)+\frac{f^{(3)}(\sigma_0)}{3!}3\sigma_1^2\sigma_2+\frac{f^{(4)}(\sigma_0)}{4!}\sigma_1^4\Big].$$

Inserting $\sigma_2,\ \sigma_3$ into this formula and keeping (6) in mind, we obtain

$$\sigma_4=\frac{\sigma_1^4}{72\lambda^3}\Big[(f'(\sigma_0))^2f^{(4)}(\sigma_0)+9(f''(\sigma_0))^3+8f'(\sigma_0)f''(\sigma_0)f^{(3)}(\sigma_0)\Big].$$

It is clear that the procedure can be continued iteratively and hence we omit any further steps for brevity. It is noted here that the complexity of this computation can be determined from (11) by a recursion. We consider this point in the next section.

4. **Bounds on Dirichlet series.** In this section we derive bounds on the terms of Dirichlet series. First we formulate the derivation of the Dirichlet series in the following theorem.

Theorem 4.1. *Let $\eta_k:=f^{(k)}(\sigma_0)$, $k=1,2,\dots$ (hence $\eta_1=\lambda$). Suppose that $\lambda\neq0$. There exist numbers $\beta_1,\beta_2,\dots,$ that depend on $\{\eta_n\}_{n=1}^{\infty}$ but not on σ_1, such that*

$$\sigma_n=\lambda\beta_n\Big(\frac{\sigma_1}{\lambda}\Big)^n,\qquad n\ge1. \tag{14}$$

Moreover, $\{\beta_n\}_{n=1}^{\infty}$ satisfy the recursion

$$\beta_1=1,$$
$$\beta_n=\frac{1}{n-1}\Big[\sum_{l=2}^{n}\frac{\eta_l\eta_1^{l-2}}{l!}\sum_{\mathbf{j}\in\widetilde{\mathbb{J}}_{n,l}}\binom{l}{j_1,j_2,\dots,j_{n-1}}\beta_1^{j_1}\beta_2^{j_2}\cdots\beta_{n-1}^{j_{n-1}}\Big],\qquad n\ge2. \tag{15}$$

Proof. It is clear that (14) is true for $n=1$ with $\beta_1=1$. We prove this theorem by induction. Suppose that (14) holds for i, an arbitrary positive integer less than or

equal to $n-1$ ($n \geq 2$) and we prove the result for n. From (11) we get

$$\begin{aligned}\sigma_n =& \frac{1}{(n-1)\lambda}\Big[\sum_{l=2}^{n}\frac{\eta_l}{l!}\sum_{\mathbf{j}\in\widetilde{\mathbb{J}}_{n,l}}\binom{l}{j_1,j_2,\ldots,j_{n-1}}\sigma_1^{j_1}\sigma_2^{j_2}\cdots\sigma_{n-1}^{j_{n-1}}\Big]\\ =& \frac{1}{(n-1)\lambda}\Big[\sum_{l=2}^{n}\frac{\eta_l}{l!}\sum_{\mathbf{j}\in\widetilde{\mathbb{J}}_{n,l}}\binom{l}{j_1,j_2,\ldots,j_{n-1}}\lambda^{j_1+j_2+\cdots+j_{n-1}}\\ &\beta_1^{j_1}\beta_2^{j_2}\cdots\beta_{n-1}^{j_{n-1}}\Big(\frac{\sigma_1}{\lambda}\Big)^{j_1+2j_2+\cdots+(n-1)j_{n-1}}\Big]\\ =& \frac{1}{(n-1)\lambda}\Big[\sum_{l=2}^{n}\frac{\eta_l}{l!}\sum_{\mathbf{j}\in\widetilde{\mathbb{J}}_{n,l}}\binom{l}{j_1,j_2,\ldots,j_{n-1}}\lambda^{l}\beta_1^{j_1}\beta_2^{j_2}\cdots\beta_{n-1}^{j_{n-1}}\Big(\frac{\sigma_1}{\lambda}\Big)^{n}\Big],\end{aligned}$$

which gives

$$\sigma_n = \lambda\Big(\frac{\sigma_1}{\lambda}\Big)^n\frac{1}{n-1}\Big[\sum_{l=2}^{n}\frac{\eta_l\eta_1^{l-2}}{l!}\sum_{\mathbf{j}\in\widetilde{\mathbb{J}}_{n,l}}\binom{l}{j_1,j_2,\ldots,j_{n-1}}\beta_1^{j_1}\beta_2^{j_2}\cdots\beta_{n-1}^{j_{n-1}}\Big]. \tag{16}$$

Therefore, (14) holds for n and β_n satisfies (15). □

Before presenting bounds on β_n, the following lemma is necessary for our analysis.

Lemma 4.2. *Using the notation* (10)*, it is true that*

$$\sum_{l=2}^{n}\sum_{\mathbf{j}\in\widetilde{\mathbb{J}}_{n,l}}\frac{1}{j_1!j_2!\cdots j_{n-1}!1^{j_1}2^{j_2}\cdots(n-1)^{j_{n-1}}} = \frac{n-1}{n}, \qquad n=2,3,\ldots. \tag{17}$$

Proof. We choose $\lambda = 1$ and $\sigma_n = 1/n$ for $n \geq 1$ in (3), i.e.,

$$y(t) = \sum_{n=1}^{\infty}\frac{1}{n}e^{nt}, \qquad t\in\mathbb{R}. \tag{18}$$

It can be verified directly that $y(t)$ of (18) satisfies $y' = e^y - 1$. This means that in (2) $f(y) = e^y - 1$ and then we obtain $\sigma_0 = 0$, $f^{(l)}(\sigma_0) = 1$ for $l \geq 1$. Under above conditions, (11) becomes

$$\begin{aligned}\frac{1}{n} &= \frac{1}{n-1}\sum_{l=2}^{n}\frac{1}{l!}\sum_{\mathbf{j}\in\widetilde{\mathbb{J}}_{n,l}}\binom{l}{j_1,j_2,\ldots,j_{n-1}}1^{j_1}\Big(\frac{1}{2}\Big)^{j_2}\cdots\Big(\frac{1}{n-1}\Big)^{j_{n-1}}\\ &= \frac{1}{n-1}\sum_{l=2}^{n}\sum_{\mathbf{j}\in\widetilde{\mathbb{J}}_{n,l}}\frac{1}{j_1!j_2!\cdots j_{n-1}!1^{j_1}2^{j_2}\cdots(n-1)^{j_{n-1}}},\end{aligned} \tag{19}$$

which proves the result (17). □

We are now in the position to present the bounds on β_n.

Theorem 4.3. *Suppose that there exists a real number $c > 0$ so that $|\eta_l| \leq c^l$, $l = 1,2,\ldots$. Then*

$$|\beta_l| \leq \frac{c^{2(l-1)}}{l}, \qquad l = 1,2,\ldots. \tag{20}$$

Proof. By Theorem 4.1, we know that $\beta_2 = \frac{1}{2}\eta_2$ and $\beta_3 = \frac{1}{12}[3\eta_2^2 + \eta_1\eta_3]$. Thus it is easy to verify that (20) holds for $l = 1, 2, 3$. We now prove this theorem by induction. Suppose (20) is true for arbitrary positive integer i ($i \le n-1$). Then (15) gives

$$\begin{aligned}|\beta_n| &\le \frac{1}{n-1}\Big[\sum_{l=2}^{n}\frac{c^{2l-2}}{l!}\sum_{\mathbf{j}\in\widetilde{\mathbb{J}}_{n,l}}\binom{l}{j_1, j_2, \ldots, j_{n-1}}|\beta_1|^{j_1}|\beta_2|^{j_2}\cdots|\beta_{n-1}|^{j_{n-1}}\Big]\\ &\le \frac{1}{n-1}\Big[\sum_{l=2}^{n}\frac{c^{2l-2}}{l!}\sum_{\mathbf{j}\in\widetilde{\mathbb{J}}_{n,l}}\frac{l!}{j_1!j_2!\cdots j_{n-1}!}\frac{c^{2[j_2+2j_3+\cdots+(n-2)j_{n-1}]}}{1^{j_1}2^{j_2}\cdots(n-1)^{j_{n-1}}}\Big].\end{aligned}$$

By (10), we have

$$j_1 + j_2 + \ldots + j_{n-1} = l,\ j_1 + 2j_2 + \ldots + (n-1)j_{n-1} = n,$$

which yields $j_2 + 2j_3 + \cdots + (n-2)j_{n-1} = n - l$. Thus

$$\begin{aligned}|\beta_n| &\le \frac{1}{n-1}\Big[\sum_{l=2}^{n}\frac{c^{2l-2}}{l!}\sum_{\mathbf{j}\in\widetilde{\mathbb{J}}_{n,l}}\frac{l!}{j_1!j_2!\cdots j_{n-1}!}\frac{c^{2(n-l)}}{1^{j_1}2^{j_2}\cdots(n-1)^{j_{n-1}}}\Big]\\ &= \frac{c^{2(n-1)}}{n-1}\Big[\sum_{l=2}^{n}\sum_{\mathbf{j}\in\widetilde{\mathbb{J}}_{n,l}}\frac{1}{j_1!j_2!\cdots j_{n-1}!1^{j_1}2^{j_2}\cdots(n-1)^{j_{n-1}}}\Big]\ \text{(using Lemma 4.2)}\\ &= \frac{c^{2(n-1)}}{n-1}\frac{n-1}{n} = \frac{c^{2(n-1)}}{n}.\end{aligned} \tag{21}$$

This means that (20) is true for $l = n$ and the proof is complete. □

5. **Global error estimate for Dirichlet series.** Under the condition of Theorem 4.3,

$$|\beta_l| \le \frac{c^{2(l-1)}}{l}, \qquad l = 1, 2, \ldots,$$

so the generating function is $g(z) := \sum_{l=1}^{\infty}\frac{c^{2(l-1)}}{l}z^l = -c^{-2}\log(1 - c^2 z)$ with the radius of convergence $R = c^{-2}$. Needless to say, for all practical purposes, we need to evaluate the finite coefficients $\sigma_2,\ \sigma_3, \ldots, \sigma_N$ and then get the truncated Dirichlet approximation which is denoted by

$$y_N(t) = \sum_{n=0}^{N}\sigma_n e^{n\lambda t}, \qquad t \in \mathbb{R}. \tag{22}$$

Subject to the aforementioned conditions, we obtain

$$\begin{aligned}|y_N(t) - y(t, y_0)| &= \Big|\sum_{n=0}^{N}\sigma_n e^{n\lambda t} - y(t, y_0)\Big| \le \lambda\sum_{n=N+1}^{\infty}|\beta_n|\Big|\frac{\sigma_1}{\lambda}e^{\lambda t}\Big|^n\\ &\le \lambda\sum_{n=N+1}^{\infty}\frac{c^{2(n-1)}}{n}\Big|\frac{\sigma_1}{\lambda}e^{\lambda t}\Big|^n,\end{aligned} \tag{23}$$

where $y(t, y_0)$ denotes the solution of (2) with the initial value $y(t_0) = y_0$.

Next we discuss the convergence of the Dirichlet expansion (22). From (23), it follows that under the condition $|\sigma_1| < |\lambda|/c^2$, a Dirichlet expansion is uniformly

convergent for all $t \geq 0$ if $\mathrm{Re}\lambda \leq 0$ or for all $t \leq \dfrac{\log(R|\frac{\lambda}{\sigma_1}|)}{\mathrm{Re}\lambda}$ if $\mathrm{Re}\lambda > 0$. Under the condition $|\sigma_1| \geq |\lambda|/c^2$, a Dirichlet expansion is uniformly convergent for all $t \geq \dfrac{\log(R|\frac{\lambda}{\sigma_1}|)}{\mathrm{Re}\lambda}$ if $\mathrm{Re}\lambda < 0$ or for all $t \leq \dfrac{\log(R|\frac{\lambda}{\sigma_1}|)}{\mathrm{Re}\lambda}$ if $\mathrm{Re}\lambda > 0$.

Remark 1. It is noted here that (23) gives a global error estimate for a global solution and it is very different from time-stepping algorithms. The Dirichlet approximation (22) yields a continuous solution that is equally valid along an interval, whose size is independent of N and which is amenable to further analytic manipulation. Meanwhile, the Dirichlet approximation does not need time stepping and just depends on t once we choose the value of N. Therefore, it is very effective and powerful in obtaining a long-time numerical solution with very small computational cost. Moreover, recall that σ_1 is a parameter, which can be tuned so as to hit the correct initial condition, similarly to shooting methods in numerical solutions. Therefore, the Dirichlet approximation delivers considerably more information than a routine time-stepping algorithm and provides another effective and different approach to solving dynamical systems of first-order ODEs.

6. **The right choice of σ_0 for Dirichlet series.** This section is concerned with the right choice of σ_0 for Dirichlet series if (2) has more than one stationary point. Here we consider the dynamical system (2), where y is a point in $\mathbb{R}$. First we refer to [16, 14] for an introduction to dynamical systems. Regarding the stationary point and its stability type, [14] presents the following elementary theorem.

Theorem 6.1. *([14]) Consider a stationary point y^* for the dynamical system $y' = f(y)$, where f and f' are continuous.*

- *If $f'(y^*) < 0$, then y^* is an asymptotically stable stationary point.*
- *If $f'(y^*) > 0$, then y^* is a repelling stationary point.*
- *If $f'(y^*) = 0$, then the derivative does not determine the stability type.*

For the definitions of stationary point, asymptotically stable stationary point and repelling stationary point, we refer the reader to [14].

From the definition of asymptotically stable stationary point, it is known that nearby solutions of (2) eventually tend to the asymptotically stable stationary point when $t \to +\infty$. Section 2 shows that σ_0 is a stationary point of the dynamical system (2) and by the scheme of Dirichlet series (3), we know that $y(t) = \sigma_0 + \sum_{n=1}^{\infty} \sigma_n e^{n\lambda t}$. If σ_0 is chosen as an asymptotically stable stationary point, so $\lambda = f'(\sigma_0) < 0$ and then $\lim_{t\to+\infty} y(t) = \sigma_0$, which coincides with the exact solution of (2). In other words, for $t \geq 0$, we should make the right choice of σ_0 such that $f'(\sigma_0) < 0$. In this case, σ_0 is an asymptotically stable stationary point and the Dirichlet approximation (3) has the same property as the exact solution of (2). Similarly, for $t \leq 0$, σ_0 should be chosen such that $f'(\sigma_0) > 0$ (a repelling stationary point).

Remark 2. The above analysis just considers $y \in \mathbb{R}$ for example. A similar result can be obtained for $y \in \mathbb{C}$. For $t \geq 0$, σ_0 should be chosen such that $\mathrm{Re}(f'(\sigma_0)) < 0$ and for $t \leq 0$, σ_0 should be chosen such that $\mathrm{Re}(f'(\sigma_0)) > 0$. Choosing the right σ_0 makes the Dirichlet approximation have the same asymptotic behaviour as the exact solution. This is a great advantage for a numerical method because it is well known that numerical methods should preserve as much as possible the qualitative behaviour of the original problem.

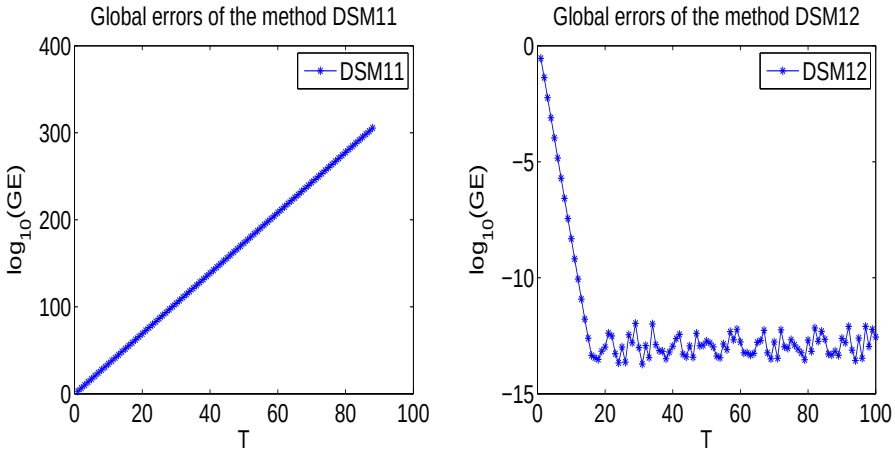

FIGURE 1. Results for Problem 1: The T against the logarithm of the global error over the integration $[0, T]$ interval for methods DSM11 (left) and DSM12 (right).

Now we use two numerical experiments to show this point and demonstrate the efficiency of the Dirichlet approximation.

Problem 1. Consider the following polynomial dynamical system

$$y' = y^2 - 2y, \qquad y(0) = 1.$$

We choose $N = 4$ in (22). There are two zeros of $y^2 - 2y = 0$ and first we choose $\sigma_0 = 2$. In this case $\lambda = 2$ and $\sigma_2 = \frac{\sigma_1^2}{2}$, $\sigma_3 = \frac{\sigma_1^3}{4}$, $\sigma_4 = \frac{\sigma_1^4}{8}$. According to the initial condition, we choose $\sigma_1 = -1.3376192231293262$ and the above method is denoted as DSM11. In the second case, $\sigma_0 = 0$ and $\lambda = -2$, $\sigma_2 = -\frac{\sigma_1^2}{2}$, $\sigma_3 = \frac{\sigma_1^3}{4}$, $\sigma_4 = -\frac{\sigma_1^4}{8}$. Consistently with the initial condition, we choose $\sigma_1 = -0.3376192231293266$ and denote this method by DSM12. We apply these two methods to solving this problem in the interval $[0, T]$, $T = i$, $i = 1, 2, \ldots, 100$ and plot the global errors $GE = |y - y_N|$ in Figure 1. In this paper, we use the result of standard ODE45 method in MATLAB with an absolute and relative tolerance equal to 10^{-12} as the true solution for all the numerical experiments. Next we solve this problem in the interval $[-T, 0]$, $T = i$, $i = 1, 2, \ldots, 100$ and the results are shown in Figure 2.

Problem 2. Consider another dynamical system

$$y' = \sin(y), \qquad y(0) = 1.$$

We also choose $N = 4$ in (22). First we choose $\sigma_0 = 0$ and $\lambda = 1$, $\sigma_2 = 0$, $\sigma_3 = -\frac{\sigma_1^3}{12}$, $\sigma_4 = 0$. Consistently with the initial condition, we choose $\sigma_1 = -3.8844837019393323$ and denote this method by DSM21. Then we choose $\sigma_0 = \pi$ and then $\lambda = -1$. Other coefficients σ_i, $i = 2, 3, 4$ are the same as those of method DSM21. From the initial condition, we choose $\sigma_1 = -2.124211319990258$ and this method is denoted as DSM22. Now we solve this problem in the interval $[0, T]$ and

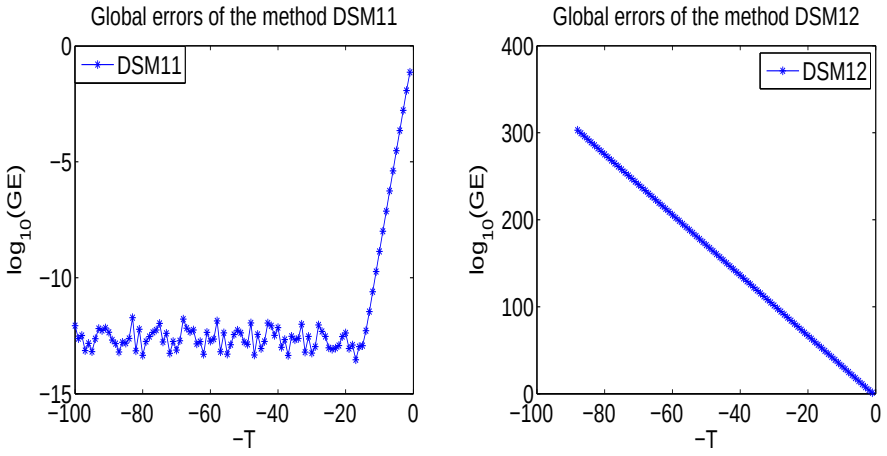

FIGURE 2. Results for Problem 1: The $-T$ against the logarithm of the global error over the integration $[-T, 0]$ interval for methods DSM11 (left) and DSM12 (right).

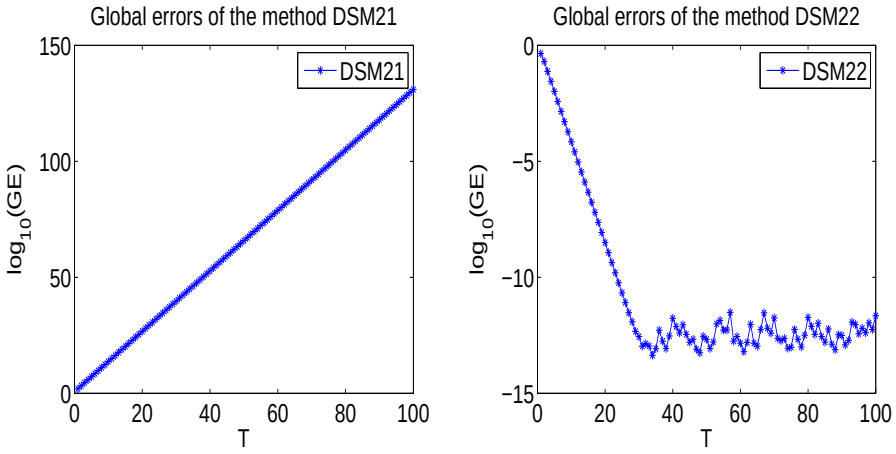

FIGURE 3. Results for Problem 2: The T against the logarithm of the global error over the integration interval $[0, T]$ for methods DSM21 (left) and DSM22 (right).

$[-T, 0]$, $T = i,\ i = 1, 2, \ldots, 100$ by the two methods. The results are given in Figure 3 and Figure 4, respectively.

It can be observed from the results of these two numerical experiments that when $t \geq 0$ and λ satisfies $\mathrm{Re}\lambda < 0$, the corresponding methods have good accuracy, whereas in this case if $\mathrm{Re}\lambda > 0$, the corresponding methods behave badly. For $t \leq 0$, the result is completely opposite. The reason for this phenomenon is that all the methods for these two problems satisfy the condition $|\sigma_1| \geq |\lambda|/c^2$. Thus, the Dirichlet method is uniformly convergent for all $t \geq \dfrac{\log(R|\frac{\lambda}{\sigma_1}|)}{\mathrm{Re}\lambda}$ if $\mathrm{Re}\lambda < 0$ or for all $t \leq \dfrac{\log(R|\frac{\lambda}{\sigma_1}|)}{\mathrm{Re}\lambda}$ if $\mathrm{Re}\lambda > 0$. Moreover, $\sigma_0 = 0$ in Problem 1 and $\sigma_0 = \pi$ in Problem

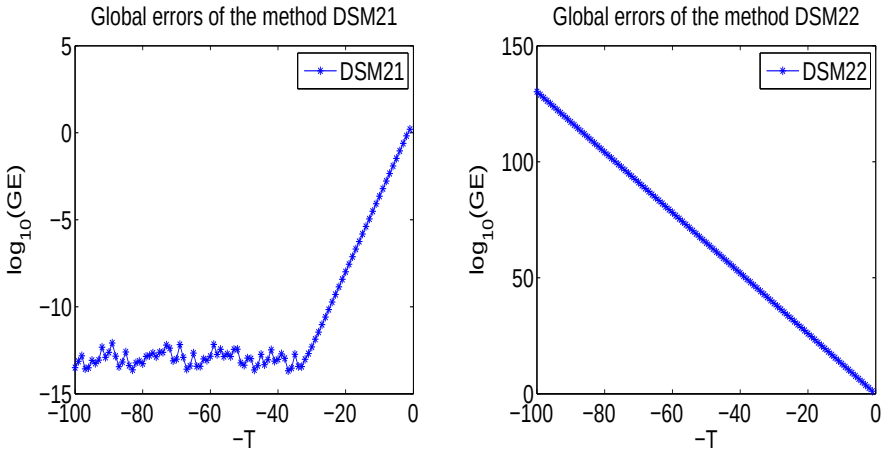

FIGURE 4. Results for Problem 2: The $-T$ against the logarithm of the global error over the integration interval $[-T, 0]$ for methods DSM21 (left) and DSM22 (right).

2 are both asymptotically stable fixed points while $\sigma_0 = 2$ in Problem 1 and $\sigma_0 = 0$ in Problem 2 are both repelling fixed points. The results of these two numerical experiments support the analysis of the convergence of Dirichlet series and show the deep connection between stable stationary points and the right choice of σ_0.

Remark 3. It is clear that there are some restrictions on applying Dirichlet series to (2) such as the requirements that the dynamical system (2) must have at least one stationary point and the first derivative of $f(y)$ evaluated at this stationary point cannot be zero. It is also required to make the right choice of σ_0 to obtain powerful Dirichlet approximations with good accuracy. However, many dynamical systems satisfy the above mentioned conditions and Dirichlet series can be considered as an effective approach to solving them as a result of the special nature of Dirichlet series. Therefore, it is very meaningful and necessary to research Dirichlet series for dynamical systems of first-order ODEs.

7. **Dirichlet series applied to multivariate dynamical systems.** In this section we discuss the application of Dirichlet series to multivariate dynamical systems. We consider the multivariate dynamical system of first-order ODEs

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}), \qquad \mathbf{y}(t_0) = \mathbf{y}_0, \tag{24}$$

where $\mathbf{y} : \mathbb{R} \to \mathbb{C}^d$ is the solution of (24) and $\mathbf{f} : \mathbb{C}^d \to \mathbb{C}^d$ is an analytic function. For this dynamical system, following (3) we consider the Dirichlet series

$$\mathbf{y}(t) = \sum_{n=0}^{\infty} e^{n\lambda t}\sigma_n, \qquad t \in \mathbb{R}, \tag{25}$$

where $\{\sigma_n\}_{n=0}^{\infty}$ are d-dimensional vectors, independent of t. First we discuss a particular case to show that if λ is a $d \times d$ matrix, generally the vectors $\{\sigma_n\}_{n=0}^{\infty}$ cannot be obtained explicitly. We consider the following particular dynamical system with

a polynomial $\mathbf{g}(\mathbf{y})$

$$\mathbf{y}' = \mathbf{g}(\mathbf{y}) = \sum_{l=0}^{s} A_l \mathbf{y}^l, \tag{26}$$

where $\mathbf{y} : \mathbb{R} \to \mathbb{C}^d$, A_l are constant $d \times d$ matrixes and $\mathbf{y}^l$ denotes $(y_1^l, y_2^l, \cdots, y_d^l)^T$. Substitution of (25) into (26) yields

$$\sum_{n=0}^{\infty} n e^{n\lambda t} \lambda \sigma_n = \sum_{l=0}^{s} A_l \Big(\sum_{k=0}^{\infty} e^{k\lambda t} \sigma_k \Big)^l. \tag{27}$$

Considering the coefficients of $e^{n\lambda t}$ with $n = 0$ and $n = 1$ in both sides of (27), we get $\mathbf{g}(\sigma_0) = \mathbf{0}$ and $\lambda = \partial \mathbf{g}(\sigma_0)/\partial \mathbf{y}$, respectively. For $n \geq 2$, if the right part of formula (27) can be reformulated as the scheme $\sum\limits_{n=0}^{\infty} e^{n\lambda t} \mathbf{x}_n$ with vectors $\mathbf{x}_n$, comparing the coefficients of $e^{n\lambda t}$ in both sides of (27) can give the expression for σ_n. However, unfortunately for $\sum\limits_{l=0}^{s} A_l \Big(\sum\limits_{k=0}^{\infty} e^{k\lambda t} \sigma_k \Big)^l$, since λ is a matrix and $\{\sigma_n\}_{n=0}^{\infty}$ are d-dimensional vectors, generally speaking

$$(e^{k_1 \lambda t} \sigma_{k_1}) \cdot (e^{k_2 \lambda t} \sigma_{k_2}) \neq (e^{k_1 \lambda t} e^{k_2 \lambda t})(\sigma_{k_1} \cdot \sigma_{k_2}), \qquad k_1, k_2 \in \mathbb{N}, \qquad k_1 \neq k_2.$$

Here $\mathbf{x} \cdot \mathbf{y}$ denotes the componentwise product of two vectors, that is, $\mathbf{x} \cdot \mathbf{y} = (x_1 y_1, x_2 y_2, \cdots, x_d y_d)^T$. Besides, in general, the matrix λ does not commute with A_l. Therefore, we cannot rewrite the right part of (27) as the scheme $\sum\limits_{n=0}^{\infty} e^{n\lambda t} \mathbf{x}_n$. This point shows that $\{\sigma_n\}_{n=2}^{\infty}$ can only be given by the formula (27) but cannot generally be expressed explicitly and independently of $e^{n\lambda t}$. However, fortunately for some particular cases, we can consider applying a new scheme of Dirichlet series to their solution.

7.1. **A new scheme of Dirichlet series for particular systems of the form (24).** In what follows we discuss a particular case for (24). Suppose that there exists $\mathbf{y}^*$ such that $\mathbf{f}(\mathbf{y}^*) = \mathbf{0}$, $\frac{\partial \mathbf{f}}{\partial \mathbf{y}}(\mathbf{y}^*) = M$, where M is a diagonalizable matrix, i.e., there exists a nonsingular matrix P such that $P^{-1} M P = \operatorname{diag}(a_1, a_2, \cdots, a_d)$. By letting $\mathbf{z} = P^{-1}\mathbf{y}$, the system (24) is equivalent to a transformed dynamical system

$$\mathbf{z}' = \mathbf{h}(\mathbf{z}) = P^{-1}\mathbf{f}(P\mathbf{z}), \qquad \mathbf{z}(t_0) = P^{-1}\mathbf{y}_0. \tag{28}$$

It is easy to check that $\mathbf{z}^* = P^{-1}\mathbf{y}^*$ is a stationary point of this transformed dynamical system and the Jacobian matrix of $\mathbf{h}(\mathbf{z})$ evaluated at $\mathbf{z}^*$ is

$$A = \operatorname{diag}(a_1, a_2, \cdots, a_d), \tag{29}$$

which means that we just need to discuss Dirichlet series for (28) and then the numerical solution of (24) can be obtained by $\mathbf{y} = P\mathbf{z}$.

It is noted that although the dynamical system (28) is of a special form, there are still many nonseparable multivariate dynamical systems fit this pattern (see the numerical experiment in Subsection 7.3 for example). Thus it is valuable to consider Dirichlet series for the multivariate dynamical system (28).

Now we seek an expansion of the solution of (28) in the following Dirichlet series

$$\mathbf{z}(t) = \sum_{n=0}^{\infty} \sum_{\mathbf{I_n} \in \mathbb{I}_{d,n}} e^{t \mathbf{I_n}^T \mathbf{a}} \alpha_{\mathbf{I_n}}, \qquad t \in \mathbb{R}, \tag{30}$$

where $\mathbf{a}=\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_d \end{pmatrix}$, $\alpha_{\mathbf{I_n}}$ is a d-dimensional vector independent of t and $\mathbb{I}_{d,n}$ is a set defined as

$$\mathbb{I}_{d,n}=\left\{\begin{pmatrix} i_1 \\ i_2 \\ \vdots \\ i_d \end{pmatrix}\in {\mathbb{Z}_+}^n:\ i_1+i_2+\ldots+i_d=n\right\}. \tag{31}$$

From this formula, it follows that

$$\mathbb{I}_{d,0}=\left\{\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\right\},\qquad \mathbb{I}_{d,1}=\left\{\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix},\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}\right\} \tag{32}$$

and we define $\alpha_{\mathbf{I_1}}$ as

$$\alpha_{(\mathbf{1},\mathbf{0},\ldots,\mathbf{0})^{\mathbf{T}}}=\begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},\ \alpha_{(\mathbf{0},\mathbf{1},\ldots,\mathbf{0})^{\mathbf{T}}}=\begin{pmatrix} 0 \\ x_2 \\ \vdots \\ 0 \end{pmatrix},\ldots,\ \alpha_{(\mathbf{0},\mathbf{0},\ldots,\mathbf{1})^{\mathbf{T}}}=\begin{pmatrix} 0 \\ 0 \\ \vdots \\ x_d \end{pmatrix} \tag{33}$$

with arbitrary numbers $x_1,x_2,\ldots,x_d$ which are independent of t. In the definition (30), we also let $\alpha_{\mathbf{I_n}}=\mathbf{0}$ if ${\mathbf{I_n}}^T\mathbf{a}=0$. For $n\geq 2$, nonzero $\alpha_{\mathbf{I_n}}$s are determined by substitution of (30) into (28). Let $\mathbf{h}_n$ denote an n-tensor related to the n-th derivative of $\mathbf{h}(\mathbf{z})$ at σ_0 with $\theta=\mathbf{z}-\sigma_0$

$$\begin{aligned}
&\mathbf{h}_0(\sigma_0) && =\mathbf{h}(\sigma_0),\\
&\mathbf{h}_1(\sigma_0)[\theta] && =\frac{\partial \mathbf{h}(\sigma_0)}{\partial \mathbf{z}}\theta,\\
&\big(\mathbf{h}_2(\sigma_0)[\theta,\theta]\big)_r && =\sum_{i=1}^{d}\sum_{j=1}^{d}\theta_i\frac{\partial^2 h_r(\sigma_0)}{\partial z_i\partial z_j}\theta_j, && r=1,2,\ldots,d,\\
&\ldots\\
&\big(\mathbf{h}_n(\sigma_0)[\theta,\ldots,\theta]\big)_r && =\sum_{i_1=1}^{d}\cdots\sum_{i_n=1}^{d}\frac{\partial^n h_r(\sigma_0)}{\partial z_{i_1}\cdots\partial z_{i_n}}\theta_{i_1}\cdots\theta_{i_n}, && r=1,2,\ldots,d.
\end{aligned} \tag{34}$$

Inserting (30) into (28) yields

$$\begin{aligned}
&\sum_{n=0}^{\infty}\sum_{\mathbf{I_n}\in\mathbb{I}_{d,n}}{\mathbf{I_n}}^T\mathbf{a}e^{t{\mathbf{I_n}}^T\mathbf{a}}\alpha_{\mathbf{I_n}}\\
&=\sum_{m=0}^{\infty}\frac{1}{m!}\mathbf{h}_m(\alpha_{\mathbf{I_0}})\Big[\sum_{n=1}^{\infty}\sum_{\mathbf{I_n}\in\mathbb{I}_{d,n}}e^{t{\mathbf{I_n}}^T\mathbf{a}}\alpha_{\mathbf{I_n}},\ldots,\sum_{n=1}^{\infty}\sum_{\mathbf{I_n}\in\mathbb{I}_{d,n}}e^{t{\mathbf{I_n}}^T\mathbf{a}}\alpha_{\mathbf{I_n}}\Big].
\end{aligned} \tag{35}$$

Considering the constant terms of this formula, we get

$$\mathbf{h}(\alpha_{\mathbf{I_0}})=\mathbf{0},$$

which shows that $\alpha_{\mathbf{I_0}}$ must be a stationary point of the dynamical system (28) and we choose $\alpha_{\mathbf{I_0}} = \mathbf{z}^*$. For $n = 1$, considering $e^{t{\mathbf{I_1}}^T\mathbf{a}}$ in both sides of (35) yields

$$\sum_{\mathbf{I_1}\in\mathbb{I}_{d,1}} {\mathbf{I_1}}^T\mathbf{a}e^{t{\mathbf{I_1}}^T\mathbf{a}}\alpha_{\mathbf{I_1}} = A\sum_{\mathbf{I_1}\in\mathbb{I}_{d,1}} e^{t{\mathbf{I_1}}^T\mathbf{a}}\alpha_{\mathbf{I_1}}. \tag{36}$$

It is easy to check that

$$\begin{aligned}&\sum_{\mathbf{I_1}\in\mathbb{I}_{d,1}} {\mathbf{I_1}}^T\mathbf{a}e^{t{\mathbf{I_1}}^T\mathbf{a}}\alpha_{\mathbf{I_1}}\\ =&a_1e^{ta_1}\begin{pmatrix}x_1\\0\\\vdots\\0\end{pmatrix} + a_2e^{ta_2}\begin{pmatrix}0\\x_2\\\vdots\\0\end{pmatrix} + \cdots + a_de^{ta_d}\begin{pmatrix}0\\0\\\vdots\\x_d\end{pmatrix} = \begin{pmatrix}a_1e^{ta_1}x_1\\a_2e^{ta_2}x_2\\\vdots\\a_de^{ta_d}x_d\end{pmatrix}\end{aligned}$$

and

$$A\sum_{\mathbf{I_1}\in\mathbb{I}_{d,1}} e^{t{\mathbf{I_1}}^T\mathbf{a}}\alpha_{\mathbf{I_1}} = \mathrm{diag}(a_1,a_2,\cdots,a_d)\begin{pmatrix}e^{ta_1}x_1\\e^{ta_2}x_2\\\vdots\\e^{ta_d}x_d\end{pmatrix} = \begin{pmatrix}a_1e^{ta_1}x_1\\a_2e^{ta_2}x_2\\\vdots\\a_de^{ta_d}x_d\end{pmatrix}.$$

Thus (36) is true and this shows that (30) is well defined for (28).

For $n \geq 2$, let

$$\mathbb{K}_{n,m} = \{(k_1,\ldots,k_m)\in\mathbb{N}^m :\ k_1+k_2+\ldots+k_m = n\}$$

with the set of natural numbers $\mathbb{N}$, and by this formula we know $n \geq m$. Now (35) can be formulated as

$$\begin{aligned}&\sum_{n=1}^{\infty}\sum_{\mathbf{I_n}\in\mathbb{I}_{d,n}} {\mathbf{I_n}}^T\mathbf{a}e^{t{\mathbf{I_n}}^T\mathbf{a}}\alpha_{\mathbf{I_n}}\\ =&\sum_{m=1}^{\infty}\sum_{n=m}^{\infty}\sum_{\mathbf{k}\in\mathbb{K}_{n,m}}\frac{1}{m!}\sum_{\mathbf{I_{k_1}}\in\mathbb{I}_{d,k_1}}\sum_{\mathbf{I_{k_2}}\in\mathbb{I}_{d,k_2}}\cdots\sum_{\mathbf{I_{k_m}}\in\mathbb{I}_{d,k_m}} e^{t{\mathbf{I_{k_1}}}^T\mathbf{a}}e^{t{\mathbf{I_{k_2}}}^T\mathbf{a}}\cdots e^{t{\mathbf{I_{k_m}}}^T\mathbf{a}}\\ &\mathbf{h}_m(\alpha_{\mathbf{I_0}})[\alpha_{\mathbf{I_{k_1}}},\alpha_{\mathbf{I_{k_2}}},\ldots,\alpha_{\mathbf{I_{k_m}}}]\\ =&\sum_{m=1}^{\infty}\sum_{n=m}^{\infty}\sum_{\mathbf{k}\in\mathbb{K}_{n,m}}\frac{1}{m!}\sum_{\mathbf{I_n}\in\mathbb{I}_{d,n}} e^{t{\mathbf{I_n}}^T\mathbf{a}}\sum_{\mathbf{I_{k_1}}+\cdots+\mathbf{I_{k_m}}=\mathbf{I_n}}\mathbf{h}_m(\alpha_{\mathbf{I_0}})[\alpha_{\mathbf{I_{k_1}}},\alpha_{\mathbf{I_{k_2}}},\ldots,\alpha_{\mathbf{I_{k_m}}}]\\ =&\sum_{n=1}^{\infty}\sum_{\mathbf{I_n}\in\mathbb{I}_{d,n}} e^{t{\mathbf{I_n}}^T\mathbf{a}}\sum_{m=1}^{n}\sum_{\mathbf{k}\in\mathbb{K}_{n,m}}\sum_{\mathbf{I_{k_1}}+\cdots+\mathbf{I_{k_m}}=\mathbf{I_n}}\frac{1}{m!}\mathbf{h}_m(\alpha_{\mathbf{I_0}})[\alpha_{\mathbf{I_{k_1}}},\alpha_{\mathbf{I_{k_2}}},\ldots,\alpha_{\mathbf{I_{k_m}}}].\end{aligned}$$

Thus

$$\begin{aligned}\alpha_{\mathbf{I_n}} =&\frac{1}{{\mathbf{I_n}}^T\mathbf{a}}\sum_{m=1}^{n}\sum_{\mathbf{k}\in\mathbb{K}_{n,m}}\sum_{\mathbf{I_{k_1}}+\cdots+\mathbf{I_{k_m}}=\mathbf{I_n}}\frac{1}{m!}\mathbf{h}_m(\alpha_{\mathbf{I_0}})[\alpha_{\mathbf{I_{k_1}}},\alpha_{\mathbf{I_{k_2}}},\ldots,\alpha_{\mathbf{I_{k_m}}}]\\ =&\frac{1}{{\mathbf{I_n}}^T\mathbf{a}}A\alpha_{\mathbf{I_n}} + \frac{1}{{\mathbf{I_n}}^T\mathbf{a}}\sum_{m=2}^{n}\sum_{\mathbf{k}\in\mathbb{K}_{n,m}}\sum_{\mathbf{I_{k_1}}+\cdots+\mathbf{I_{k_m}}=\mathbf{I_n}}\frac{1}{m!}\mathbf{h}_m(\alpha_{\mathbf{I_0}})[\alpha_{\mathbf{I_{k_1}}},\alpha_{\mathbf{I_{k_2}}},\ldots,\alpha_{\mathbf{I_{k_m}}}].\end{aligned}$$

Therefore, under the condition that

$${\mathbf{I_n}}^T\mathbf{a} - a_k \neq 0, \qquad k = 1,2,\ldots,d, \tag{37}$$

we have

$$\alpha_{\mathbf{I_n}} = \big(\mathbf{I_n}^T \mathbf{a} I - A\big)^{-1} \sum_{m=2}^{n} \sum_{\mathbf{k}\in\mathbb{K}_{n,m}} \sum_{\mathbf{I_{k_1}}+\cdots+\mathbf{I_{k_m}}=\mathbf{I_n}} \frac{1}{m!} \mathbf{h}_m(\alpha_{\mathbf{I_0}})[\alpha_{\mathbf{I_{k_1}}}, \alpha_{\mathbf{I_{k_2}}}, \ldots, \alpha_{\mathbf{I_{k_m}}}], \ n \geq 2, \tag{38}$$

where I is the identity matrix of the same order as A. It should be noted here that (38) is an explicit formula for nonzero $\alpha_{\mathbf{I_n}}$ and by which, we can derive explicitly all the terms of Dirichlet series (30). And in practical applications, we usually evaluate finite number of terms $\alpha_{\mathbf{I_n}}$, $n = 2, 3, \ldots, N$. Thus the condition (37) is just needed for $n = 2, 3, \ldots, N$ and this is possible and realizable for many practical problems. Moreover, recall that $x_1, x_2, \ldots, x_d$ in (33) are parameters, which can be fine tuned so as to hit the correct initial condition in (28).

Now we just present $\alpha_{\mathbf{I_2}}$ with $d = 2$ as an example. Formula (31) gives

$$\mathbb{I}_{2,2} = \left\{ \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right\}.$$

Thus, from (38) we have

$$\begin{aligned}
\alpha_{(\mathbf{2},\mathbf{0})^\intercal} &= \big(2a_1 I - A\big)^{-1} \frac{1}{2} \mathbf{h}_2(\alpha_{\mathbf{I_0}})[\alpha_{(\mathbf{1},\mathbf{0})^\intercal}, \alpha_{(\mathbf{1},\mathbf{0})^\intercal}], \\
\alpha_{(\mathbf{1},\mathbf{1})^\intercal} &= \big((a_1 + a_2) I - A\big)^{-1} \frac{1}{2} \big(\mathbf{h}_2(\alpha_{\mathbf{I_0}})[\alpha_{(\mathbf{1},\mathbf{0})^\intercal}, \alpha_{(\mathbf{0},\mathbf{1})^\intercal}] + \mathbf{h}_2(\alpha_{\mathbf{I_0}})[\alpha_{(\mathbf{0},\mathbf{1})^\intercal}, \alpha_{(\mathbf{1},\mathbf{0})^\intercal}]\big), \\
\alpha_{(\mathbf{0},\mathbf{2})^\intercal} &= \big(2a_2 I - A\big)^{-1} \frac{1}{2} \mathbf{h}_2(\alpha_{\mathbf{I_0}})[\alpha_{(\mathbf{0},\mathbf{1})^\intercal}, \alpha_{(\mathbf{0},\mathbf{1})^\intercal}].
\end{aligned}$$

7.2. **Some results for the new Dirichlet series.** In this paper, the Euclidean norm for a matrix or a vector is denoted by $\|\cdot\|$. Letting $\theta_l = \left\| \dfrac{\partial^l \mathbf{h}(\alpha_{\mathbf{I_0}})}{\partial \mathbf{z}^l} \right\|$, we have the following result about the bounds on the terms $\alpha_{\mathbf{I_n}}$,

Theorem 7.1. *Suppose that for any $\mathbf{I_1} \in \mathbb{I}_{d,1}$, $\|\alpha_{\mathbf{I_1}}\| \leq C$, then there exist real numbers $\delta_{\mathbf{I_n}}$ that depend on $\{\theta_n\}_{n=1}^{\infty}$ but not on C, such that*

$$\|\alpha_{\mathbf{I_n}}\| \leq \delta_{\mathbf{I_n}} C^n, \qquad \mathbf{I_n} \in \mathbb{I}_{d,n}, \qquad n \geq 1. \tag{39}$$

Here $\{\delta_{\mathbf{I_n}}\}_{n=1}^{\infty}$ satisfy the recursion

$$\begin{aligned}
&\delta_{\mathbf{I_1}} \equiv 1, \ \mathbf{I_1} \in \mathbb{I}_{d,1}, \\
&\delta_{\mathbf{I_n}} = \left\| \big(\mathbf{I_n}^T \mathbf{a} I - A\big)^{-1} \right\| \sum_{m=2}^{n} \sum_{\mathbf{k}\in\mathbb{K}_{n,m}} \sum_{\mathbf{I_{k_1}}+\cdots+\mathbf{I_{k_m}}=\mathbf{I_n}} \frac{\theta_m}{m!} \delta_{\mathbf{I_{k_1}}} \delta_{\mathbf{I_{k_2}}} \cdots \delta_{\mathbf{I_{k_m}}}, \qquad n \geq 2.
\end{aligned} \tag{40}$$

This theorem can easily be proved by induction and so we skip its proof here.

Insofar as other aspects of Dirichlet series (30) (e.g., global error and the right choice of σ_0), we just present them by the following two theorems without proofs, because they are similar to those of Dirichlet series (3) for scalar dynamical systems.

Theorem 7.2. *Denoting the truncation of the new Dirichlet series (30) by*

$$\mathbf{z}_N(t) = \sum_{n=0}^{N} \sum_{\mathbf{I_n}\in\mathbb{I}_{d,n}} e^{t\mathbf{I_n}^T \mathbf{a}} \alpha_{\mathbf{I_n}}, \qquad t \in \mathbb{R}, \tag{41}$$

the global error for this approximation is

$$\|\mathbf{z}_N(t) - \mathbf{z}(t, z(t_0))\| \leq \sum_{n=N+1}^{\infty} \sum_{\mathbf{I_n} \in \mathbb{I}_{d,n}} \left| e^{t\mathbf{I_n}^T \mathbf{a}} \right| \|\alpha_{\mathbf{I_n}}\|,$$

where $\|\alpha_{\mathbf{I_n}}\|$ *is given in Theorem 7.1 and* $\mathbf{z}(t, z(t_0))$ *denotes the solution of* (28).

Theorem 7.3. *The right choice of* σ_0 *in the new Dirichlet approximation* (41) *requires that*

- *for* $t \geq 0$ σ_0 *should be chosen such that* $Re(a_k) \leq 0$*;*
- *for* $t \leq 0$ σ_0 *should be chosen such that* $Re(a_k) \geq 0$*,*

where a_k *is defined in* (29) *and* $k = 1, 2, \ldots, d$.

7.3. **A numerical example for multivariate dynamical systems.** Consider the following multivariate dynamical system

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}) = \begin{pmatrix} -y_1^2 + \frac{3}{2}y_2 - \frac{1}{2} \\ \frac{1}{2} - \frac{1}{2}y_2 \end{pmatrix}, \qquad \mathbf{y}(0) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

The Jacobian matrix of $\mathbf{f}(\mathbf{y})$ evaluated at the stationary point $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is a diagonalizable matrix $\begin{pmatrix} -2 & \frac{3}{2} \\ 0 & -\frac{1}{2} \end{pmatrix}$. Let $\mathbf{z} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \mathbf{y}$ and then this problem can be transformed into

$$\mathbf{z}' = \mathbf{h}(\mathbf{z}) = \begin{pmatrix} -(z_1 + z_2)^2 + 2z_2 - 1 \\ \frac{1}{2} - \frac{1}{2}z_2 \end{pmatrix}, \qquad \mathbf{z}(0) = \begin{pmatrix} -1 \\ 2 \end{pmatrix}. \tag{42}$$

It is easy to check that $\mathbf{z_0} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is a stationary point of (42) and $\mathrm{diag}(-2, -\frac{1}{2})$ is the Jacobian matrix of $\mathbf{h}(\mathbf{z})$ evaluated at $\mathbf{z_0}$.

We choose $N = 2$ in (41) and the corresponding method is

$$\mathbf{z}_{num} = \mathbf{z_0} + \begin{pmatrix} x_1 e^{-2t} \\ x_2 e^{-\frac{1}{2}t} \end{pmatrix} + e^{-4t} \begin{pmatrix} \frac{1}{2}x_1^2 \\ 0 \end{pmatrix} + e^{-\frac{5}{2}t} \begin{pmatrix} 4x_1 x_2 \\ 0 \end{pmatrix} + e^{-t} \begin{pmatrix} -x_2^2 \\ 0 \end{pmatrix},$$

where x_1, x_2 are arbitrary numbers in (33). By the initial condition in (42), we choose $x_1 = -10$, $x_2 = 1$ and then denote the method as DSM3. We apply this method to solving (42) in the interval $[0, t]$, $t = 1, 2, \ldots, 100$ and plot the global errors $\left(GE = \left\| \mathbf{y} - \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \mathbf{z}_{num} \right\|\right)$ in Figure 5.

7.4. **A very special case of** (28)**.** Now we discuss a very special case of the ODE (28). Suppose that for matrix A in (29), $a_1 = a_2 = \ldots = a_d = \alpha$ with $\alpha \in \mathbb{C}$. In this case, we can use a simple variation on the theme of Dirichlet series (25), choosing λ in (25) as $\lambda = \alpha$. Letting $\nu = \sum_{n=1}^{\infty} e^{n\lambda t}\sigma_n$ and inserting (25) into (28) yield

$$\sum_{n=0}^{\infty} n\lambda e^{n\lambda t}\sigma_n = \mathbf{h}(\mathbf{z}) = \sum_{l=0}^{\infty} \frac{1}{l!}\mathbf{h}_l(\sigma_0)[\nu, \ldots, \nu] = \sum_{n=0}^{\infty} e^{n\lambda t} \sum_{l=0}^{\infty} \frac{1}{l!} \sum_{\mathbf{a} \in \mathbb{K}_{n,l}} \mathbf{h}_l(\sigma_0)[\sigma_{\mathbf{a_1}}, \ldots, \sigma_{\mathbf{a_l}}]. \tag{43}$$

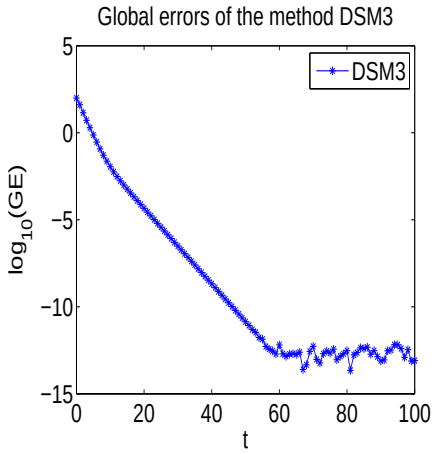

FIGURE 5. Results for DSM3: The t against the logarithm of the global error over the integration $[0, t]$.

Letting $n = 0$ in (43) gives $\mathbf{h}(\sigma_0) = \mathbf{0}$ and so we choose $\sigma_0 = \mathbf{z}^*$. Similarly, we get $\lambda = \alpha$ by letting $n = 1$. For $n \geq 2$, from (43), it follows that

$$\begin{aligned}\sigma_n &= \frac{1}{n\lambda}\sum_{l=1}^{n}\frac{1}{l!}\sum_{\mathbf{a}\in\mathbb{K}_{n,l}}\mathbf{h}_l(\sigma_0)[\sigma_{\mathbf{a_1}},\ldots,\sigma_{\mathbf{a_l}}]\\ &= \frac{1}{n\lambda}\lambda I\sigma_n + \frac{1}{n\lambda}\sum_{l=2}^{n}\frac{1}{l!}\sum_{\mathbf{a}\in\mathbb{K}_{n,l}}\mathbf{h}_l(\sigma_0)[\sigma_{\mathbf{a_1}},\ldots,\sigma_{\mathbf{a_l}}].\end{aligned}$$

Thus

$$\sigma_n = \frac{1}{(n-1)\lambda}\sum_{l=2}^{n}\frac{1}{l!}\sum_{\mathbf{a}\in\mathbb{K}_{n,l}}\mathbf{h}_l(\sigma_0)[\sigma_{\mathbf{a_1}},\ldots,\sigma_{\mathbf{a_l}}], \qquad n = 2,3,\ldots. \tag{44}$$

By this pattern, each σ_n of $n \geq 2$ can be obtained through an iterated procedure. Using the notation in (10) and by (44), it is easy to check that

$$\begin{aligned}\|\sigma_n\| &\leq \frac{1}{(n-1)|\lambda|}\sum_{l=2}^{n}\frac{1}{l!}\sum_{\mathbf{a}\in\mathbb{K}_{n,l}}\|\mathbf{h}_l(\sigma_0)[\sigma_{\mathbf{a_1}},\ldots,\sigma_{\mathbf{a_l}}]\|\\ &\leq \frac{1}{(n-1)|\lambda|}\sum_{l=2}^{n}\frac{\theta_l}{l!}\sum_{\mathbf{a}\in\mathbb{K}_{n,l}}\|\sigma_{a_1}\|\,\|\sigma_{a_2}\|\cdots\|\sigma_{a_l}\|\\ &\leq \frac{1}{(n-1)|\lambda|}\sum_{l=2}^{n}\frac{\theta_l}{l!}\sum_{\mathbf{j}\in\mathbb{JJ}_{n,l}}\binom{l}{j_1,j_2,\ldots,j_{n-1}}\|\sigma_1\|^{j_1}\|\sigma_2\|^{j_2}\cdots\|\sigma_{n-1}\|^{j_{n-1}}\end{aligned} \tag{45}$$

with $\theta_l = \left\|\dfrac{\partial^l\mathbf{h}(\sigma_0)}{\partial\mathbf{z}^l}\right\|$. With regard to the bounds on the terms σ_n, we have the following result.

Theorem 7.4. *There exist numbers $\gamma_1, \gamma_2, \ldots$, that depend on $\{\theta_n\}_{n=1}^{\infty}$ but not on $\|\sigma_1\|$, such that*

$$\|\sigma_n\| \le |\lambda|\gamma_n \big(\frac{\|\sigma_1\|}{|\lambda|}\big)^n, \qquad n \ge 1. \tag{46}$$

Here $\{\gamma_n\}_{n=1}^{\infty}$ satisfy the recursion

$$\gamma_1 = 1,$$

$$\gamma_n = \frac{1}{n-1}\Big[\sum_{l=2}^{n} \frac{\theta_l \theta_1^{l-2}}{l!} \sum_{\mathbf{j}\in \mathbb{JJ}_{n,l}} \binom{l}{j_1, j_2, \ldots, j_{n-1}} \gamma_1^{j_1}\gamma_2^{j_2}\cdots\gamma_{n-1}^{j_{n-1}}\Big], \qquad n \ge 2. \tag{47}$$

Moreover, suppose that a real number $c > 0$ exists so that $\theta_l \le c^l$, $l = 1, 2, \ldots$. Then

$$\gamma_l \le \frac{c^{2(l-1)}}{l}, \qquad l = 1, 2, \ldots. \tag{48}$$

The proof of this theorem is similar to that given in Section 4 and so we skip it here. Other discussions including global error and the right choice of σ_0 are similar to those of Dirichlet series (3) for scalar systems. Therefore we do not discuss them again.

8. **Conclusions.** In this paper we first focus our attention on applying Dirichlet series to the solution of the scalar dynamical system (2). This approach to solving this dynamical system is different from time-stepping algorithms and the analysis developed in this paper gains new insight into this method by presenting bounds and showing the global errors for this method. We also discuss the convergence of the method and consider the right choice of σ_0. Then we venture further into the topic of Dirichlet series, applying them to multivariate dynamical systems. We define a new Dirichlet expansion for a special multivariate dynamical system and present some results about this new method. The great advantage of Dirichlet series is that it does not need time stepping and yields a continuous solution that depends only on t, which is significant for obtaining long-time numerical solution. Three numerical experiments are carried out in this paper to demonstrate the efficiency and robustness of the Dirichlet approximation for scalar dynamical systems and multivariate dynamical systems.

Acknowledgements. The authors are grateful to Professor Christian Lubich and Professor Xinyuan Wu for their careful reading of the manuscript and for their helpful comments, which considerably improve its presentation.

REFERENCES

[1] J. C. Butcher, "Numerical Methods for Ordinary Differential Equations", 2nd ed., John Wiley and Sons, Ltd 2008.

[2] Y. Fang, Y. Song and X. Wu, *New embedded pairs of explicit Runge-Kutta methods with FSAL properties adapted to the numerical integration of oscillatory problems*, Phys. Lett. A, **372** (2008), 6551–6559.

[3] A. B. González, P. Martín and J. M. Farto, *A new family of Runge-Kutta type methods for the numerical integration of perturbed oscillators*, Numer. Math., **82** (1999), 635–646.

[4] E. Hairer, C. Lubich and G. Wanner, *Geometric numerical integration illustrated by the Störmer-Verlet method*, Acta Numer., **12** (2003), 399–450.

[5] E. Hairer, S. P. Nørsett and G. Wanner, "Solving Ordinary Differential Equations I: Nonstiff Problems", Springer-Verlag, Berlin, 1993.

[6] G.H. Hardy and M. Riesz, "The general theory of Dirichlet series", Cambridge Tracts in Mathematics and Mathematical Physics, Cambridge Univ. Press, London, 1915.

[7] M. Hochbruck and A. Ostermann, *Explicit exponential Runge-Kutta methods for semilineal parabolic problems*, SIAM J. Numer. Anal., **43** (2005), 1069–1090.
[8] A. Iserles, "A First Course in the Numerical Analysis of Differential Equations", 2nd ed., Cambridge University Press, Cambridge, 2008.
[9] A. Iserles, G.P. Ramaswami and M. Sofroniou, *Runge-Kutta methods for quadratic ordinary differential equations*, BIT, **38** (1998), 315–346.
[10] A. Iserles and G. Söderlind, *Global bounds on numerical error for ordinary differential equations*, J. Complexity, **9** (1993), 97–112.
[11] A. Iserles and A. Zanna, *Preserving algebraic invariants with Runge-Kutta methods*, J. Comp. Appl. Maths., **125** (2000), 69–81.
[12] S. Mandelbrojt, "Dirichlet series: principles and methods", D. Reidel Publishing Company, Dordrecht, Holland, 1972.
[13] L. Perko, "Differential Equations and Dynamical Systems", 3rd ed., Springer-Verlag, New York, 2001.
[14] R.C. Robinson, "An Introduction to Dynamical Systems: Countinuous and Discrete", Pearson Prentice Hall, New Jersey, 2004.
[15] J.M. Sanz-Serna, *Runge-Kutta schems for Hamiltonian systems*, BIT, **28** (1988), 877–883.
[16] F. Verhulst, "Nonlinear Differential Equations and Dynamical Systems", Springer-Verlag, New York, 1990.

E-mail address: wangbinmaths@gmail.com
E-mail address: A.Iserles@damtp.cam.ac.uk