# STABILITY AND SPECTRAL CONVERGENCE OF FOURIER METHOD FOR NONLINEAR PROBLEMS. ON THE SHORTCOMINGS OF THE 2/3 DE-ALIASING METHOD

## CLAUDE BARDOS AND EITAN TADMOR

ABSTRACT. The high-order accuracy of Fourier method makes it the method of choice in many large scale simulations. We discuss here the stability of Fourier method for nonlinear evolution problems, focusing on the two prototypical cases of the inviscid Burgers' equation and the multi-dimensional incompressible Euler equations. The Fourier method for such problems with quadratic nonlinearities comes in two main flavors. One is the spectral Fourier method. The other is the 2/3 pseudo-spectral Fourier method, where one removes the highest 1/3 portion of the spectrum; this is often the method of choice to maintain the balance of quadratic energy and avoid aliasing errors.

Two main themes are discussed in this paper. First, we prove that as long as the underlying exact solution has a minimal  $C^{1+\alpha}$  spatial regularity, then both the spectral and the 2/3 pseudo-spectral Fourier methods are stable. Consequently, we prove their spectral convergence for smooth solutions of the inviscid Burgers equation and the incompressible Euler equations. On the other hand, we prove that after a critical time at which the underlying solution lacks sufficient smoothness, then both the spectral and the 2/3 pseudo-spectral Fourier methods exhibit nonlinear instabilities which are realized through spurious oscillations. In particular, after shock formation in inviscid Burgers' equation, the total variation of bounded (pseudo-) spectral Fourier solutions must increase with the number of increasing modes and we stipulate the analogous situation occurs with the 3D incompressible Euler equations: the limiting Fourier solution is shown to enforce  $L^2$ -energy conservation, and the contrast with energy dissipating Onsager solutions is reflected through spurious oscillations.

## Contents

1.	Introduction	2
2.	Fourier method for linear equations — weak instability for $L^2$ -data	5
3.	Fourier method for Burgers equation: convergence for smooth solutions	11
4.	Fourier method for Burgers equation: instability for weak solutions	16
5.	Fourier method for Euler equations: convergence for smooth solutions	18
6.	Fourier method for Euler equations: failure of convergence for weak solutions?	21
7.	The spectral viscosity method: nonlinear stability and spectral convergence	23
8.	Beyond quadratic nonlinearities: 1D isentropic equations	25
Re	References	

Date: August 24, 2013.

<sup>1991</sup> Mathematics Subject Classification. 65M12,65M70,35L65,35Q31.

Key words and phrases. Spectral and pseudo-spectral methods, aliasing, stability, spectral convergence, Burgers equation, incompressible Euler equations.

Acknowledgment. E. T. Research was supported in part by NSF grants DMS10-08397, RNMS11-07444 (KI-Net) and ONR grant #N00014-1210318.

## C. BARDOS AND E. TADMOR

## 1. INTRODUCTION

Spectral methods are often the methods of choice when high-resolution solvers are sought for nonlinear time-dependent problems. Here, we are concerned with the stability and convergence of Fourier method for PDEs with quadratic nonlinearities: we focus our attention on the prototypical Cauchy problems for the inviscid Burgers' equation and the incompressible Euler equations.

The Fourier methods for problems involving quadratic nonlinearities come in two main flavors: the spectral Fourier method and the 2/3 smoothing of pseudo-spectral Fourier method. The spectral Fourier method is realized in terms of N-degree Fourier expansions,  $\mathbf{u}_N(\mathbf{x},t) = \sum_{|\mathbf{k}| < N} \hat{\mathbf{u}}_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{x}}$ , where  $\hat{\mathbf{u}}_{\mathbf{k}}(t)$  are the Fourier moments of  $\mathbf{u}(\mathbf{x},t)$ 

$$\widehat{\mathbf{u}}_{\mathbf{k}}(t) = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \mathbf{u}(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} d\mathbf{x}, \quad \mathbf{k} := (k_1, \dots, k_d) \in \mathbb{Z}^d.$$

The computation of these moments in nonlinear problems is carried out by convolutions. These can be avoided when the  $\hat{\mathbf{u}}_{\mathbf{k}}$ 's are replaced by the discrete Fourier coefficients, sampled at the  $(2N+1)^d$  equally spaced grid points

$$\widetilde{\mathbf{u}}_{\mathbf{k}}(t) = \left(\frac{1}{2N+1}\right)^d \sum_{\mathbf{x}_{\boldsymbol{\nu}} \in \mathbb{T}_{\#}^d} \mathbf{u}(\mathbf{x}_{\boldsymbol{\nu}}, t) e^{-i\mathbf{k} \cdot \mathbf{x}_{\boldsymbol{\nu}}}, \qquad \mathbf{x}_{\boldsymbol{\nu}} = \frac{2\pi\boldsymbol{\nu}}{2N+1},$$

where  $\mathbb{T}^d_{\#}$  is the discrete torus,

$$\mathbb{T}^d_{\#} := \left\{ \mathbf{x}_{\boldsymbol{\nu}} \mid \mathbf{x}_{\boldsymbol{\nu}} = \frac{2\pi\boldsymbol{\nu}}{2N+1}, \quad \boldsymbol{\nu} = (\nu_1, \dots, \nu_d), \ 0 \le \nu_j \le 2N \right\}.$$

The pseudo-spectral Fourier method is realized in terms of the corresponding expansion,  $\mathbf{u}_N(\mathbf{x},t) = \sum_{|\mathbf{k}| \leq N} \widetilde{\mathbf{u}}_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{x}}$ . Here, we have the advantage that nonlinearities are computed as exact pointwise quantities at the grid points  $\{\mathbf{x}_{\boldsymbol{\nu}}\}_{\boldsymbol{\nu}}$ , but new aliasing errors are introduced. To avoid aliasing errors and their potential instabilities, high mode smoothing is implemented, which results in the so-called 2/3-smoothing of pseudo-spectral Fourier method: it is realized in terms of the 2N/3-degree expansion,  $\mathbf{u}_N(\mathbf{x},t) = \sum_{|\mathbf{k}| \leq 2N/3} \sigma_{\mathbf{k}} \widetilde{\mathbf{u}}_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{x}}$ . This is the spectral method of choice in many time-dependent problems with quadratic nonlinearities.

To put our discussion into perspective we begin, in section 2, by recalling the linear setup of standard transport equation. The spectral Fourier method is  $L^2$ -stable. But the pseudospectral Fourier method is not [GHT94]: it is only weakly stable, due to amplification of aliasing errors when the underlying solution lacks sufficient smoothness. Strong  $L^2$ stability is regained with the 2/3-smoothing of pseudo-spectral Fourier method, [Tad87]; in the linear setup, the de-aliasing in the 2/3-method introduces sufficient smoothness to maintain convergence. This is one of the main two themes of our results on nonlinear problems: sufficient smoothness guarantees stability and hence spectral convergence. In section 3 we explore this issue in the context of inviscid Burgers equations, proving that as long as the solution of the inviscid Burgers equation remains smooth,  $u(\cdot, t) \in C_x^{1+\alpha}$ , then both the spectral and the 2/3-pseudo-spectral Fourier approximations,  $u_N(\cdot, t)$ , converge to the exact solution. Moreover, they enjoy spectral convergence rate, namely, the convergence rate grows with the increasing smoothness of  $u(\cdot, t)$ ,

$$\int |u_N(x,t) - u(x,t)|^2 dx$$

ON THE STABILITY AND INSTABILITIES OF THE FOURIER METHOD

$$\lesssim e^{\int_0^t \|u_x(\cdot,\tau)\|_{L^{\infty}} d\tau} \cdot \left( N^{-2s} \|u(\cdot,0)\|_{H^s}^2 + N^{\frac{3}{2}-s} \max_{\tau \le t} \|u(\cdot,\tau)\|_{H^s} \right), \quad s > \frac{3}{2}.$$

A similar statement of spectral convergence holds for the spectral and 2/3 pseudo-spectral Fourier approximations  $\mathbf{u}_N$  of the incompressible Euler equations: in section 5 we prove that as long as  $\mathbf{u}(\cdot, t)$  remains sufficiently smooth solution of the *d*-dimensional Euler equations,  $\mathbf{u}(\cdot, t) \in C_x^{1+\alpha}$ , then

$$\begin{aligned} \|\mathbf{u}_{N}(\cdot,t) - \mathbf{u}(\cdot,t)\|_{L^{2}}^{2} \\ \lesssim e^{2} \int_{0}^{t} \|\nabla_{\mathbf{x}}\mathbf{u}(\cdot,\tau)\|_{L^{\infty}} d\tau \\ & \quad \left(N^{-2s} \|\mathbf{u}(\cdot,0)\|_{H^{s}}^{2} + N^{\frac{d}{2}+1-s} \max_{\tau \leq t} \|\mathbf{u}(\cdot,\tau)\|_{H^{s}}\right), \ s > \frac{d}{2} + 1. \end{aligned}$$

These results support the superiority of spectral methods for problems with smooth solutions. When dealing with solutions which lack smoothness, however, both the spectral and 2/3 pseudo-spectral Fourier methods suffer nonlinear instabilities. This is the other main theme of the paper, explored in the context of the inviscid Burgers equation and the incompressible Euler equations in the respective sections 4 and 6. In particular, we prove that after shock formation, the spectral and 2/3 pseudo-spectral bounded approximations of the inviscid Burgers solution must produce spurious oscillations as their total variation must increase,  $||u_N(\cdot,t)||_{TV} \gtrsim \sqrt[4]{N}$ . This is deduced by contradiction: in theorem 4.1 below we prove, using compensated compactness arguments, that an  $L^2$ -weak limit of slowly growing TV Fourier solutions,  $\overline{u} = w \lim u_N$ , must be an  $L^2$ -energy conservative solution, which cannot hold once shocks are formed.

A similar scenario arises with the Euler solutions where the spectral and the (2/3) pseudo-)spectral approximations of Euler equations enforce conservation of the  $L^2$ -energy. Although there is no known energy dissipation-based selection principle to identify a unique solution of Euler equations within the class of "rough" data (similar to the entropy dissipation selection principle for Burgers' equations), nevertheless we argue that the  $L^2$ -energy conservation of the (pseudo-)spectral approximations may be responsible to their unstable behavior. While  $L^2$ -energy conservation holds for weak solutions with a minimal degree of 1/3-order of smoothness (Onsager's conjecture proved in [Ey94, CET94, BT10]), there are experimental and numerical evidence for the other part of Onsager's conjecture that anomalous dissipation of energy shows up for "physical-turbulent"  $L^2$ -solutions of Euler equations [Co07]. Whether this observed anomalous dissipation of energy should be due to spontaneous appearance of singularities in smooth solutions of the Euler equation or to the fact that physical initial data may be rough is a completely open problem. However after several preliminary breakthrough [Sc93] and [Sh97] the following fact are now well established. Indeed, there are infinitely many initial data (which of course are not regular) leading to infinitely many weak Euler solutions with energy loss [DeLS12]. In particular there are energy decaying solutions which for almost every time belong to the critical regularity  $C^{\frac{1}{3}-\epsilon}$  [Buck13]. Thus, if the numerical method captures such "rough" solutions then the "unphysical" conservation of energy which is enforced at the spectral level has to vanish at the limit, leading to spurious oscillations.

We close this paper with two complementary results. The nonlinear instability results in sections 4 and 6 emphasize the competition between spectral convergence for smooth solutions vs. nonlinear instabilities for problems which lack sufficient smoothness. In section 7 we discuss the class of *spectral viscosity* (SV) methods which entertain both — spectral convergence and nonlinear stability, [Tad89, Tad93b, KK00, GP03, SS07, AX09]. This is achieved by adding a judicious amount of spectral viscosity at the high-portion of the spectrum without sacrificing the spectral accuracy at the lower portion of the spectrum. Finally, we note that the above stability and instability results are not necessarily restricted to quadratic nonlinearities: in section 8 we prove the stability of Fourier method for smooth solutions of the nonlinear isentropic equations.

1.1. Spectral convergence. Expressed in terms of the Fourier coefficients,  $\widehat{w}(\mathbf{k})$ , the spectral Fourier projection  $P_N[w](\mathbf{x})$  of  $w \in L^1[\mathbb{T}^d]$  is given by

$$P_N[w](\mathbf{x}) = \sum_{|\mathbf{k}| \le N} \widehat{w}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}, \qquad \widehat{w}(\mathbf{k}) := \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} w(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} d\mathbf{x}, \quad \mathbf{k} := (k_1, \dots, k_d) \in \mathbb{Z}^d.$$

The convergence rate of the *truncation error*,

(1.1) 
$$(I - P_N)[w](\mathbf{x}) := \sum_{|\mathbf{k}| \ge N} \widehat{w}(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}}$$

is as rapid as the global smoothness of w permits (and observe that the degree of smoothness is allowed to be negative),

$$||(I - P_N)[w]||_{\dot{H}^r} \le N^{r-s} ||w||_{\dot{H}^s}, \qquad s > r \in \mathbb{R};$$

in particular,

(1.2) 
$$\max_{\mathbf{x}} |(I - P_N)[w](\mathbf{x})| \lesssim N^{\frac{d}{2} - s} ||w||_{H^s}, \qquad s > \frac{d}{2}.$$

These are statements of spectral convergence rate: the smoother w is, the faster is the convergence rate of  $(I - P_N)[w] \rightarrow 0$ . In practice, one recovers exponential convergence which characterizes analytic regularity or at least root-exponential rate for typical compactly supported Gevrey-regular data, [Tad07].

1.2. Aliasing. Set  $h := \frac{2\pi}{2N+1}$  as a discrete spacing. If we replace the integrals with quadrature based on sampling w at the  $(2N+!)^d$  equi-spaced points,  $\mathbf{x}_{\boldsymbol{\nu}} := \boldsymbol{\nu}h, \ \boldsymbol{\nu} := (\nu_1, \ldots, \nu_d) \in \{0, 2N\}^d$ , we obtain the *pseudo-spectral* Fourier projection,

$$\psi_N[w](\mathbf{x}) = \sum_{|\mathbf{k}| \le N} \widetilde{w}(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}}, \qquad \widetilde{w}(\mathbf{k}) := \left(\frac{h}{2\pi}\right)^d \sum_{\mathbf{x}_{\boldsymbol{\nu}} \in \mathbb{T}^d_{\#}} w(\mathbf{x}_{\boldsymbol{\nu}}) e^{-i\mathbf{k} \cdot \mathbf{x}_{\boldsymbol{\nu}}}, \quad |\mathbf{k}| \le N.$$

Here,  $\widetilde{w}(\mathbf{k})$ , are the discrete Fourier coefficients<sup>1</sup>. The mapping  $w \mapsto \psi_N[w]$  is a projection:  $\psi_N[w](\mathbf{x})$  is the unique N-degree trigonometric interpolant of w at the (2N+1)-gridpoints,  $\psi_N[w](\mathbf{x}_{\boldsymbol{\nu}}) = w(\mathbf{x}_{\boldsymbol{\nu}}), \ |\boldsymbol{\nu}| \leq 2N$ . The dual statement of the last equalities is the Poisson summation formula, which determines the discrete  $\widetilde{w}(\mathbf{k})$ 's in terms of the exact Fourier coefficients,  $\widehat{w}(\mathbf{k})$ 's,

$$\widetilde{w}(\mathbf{k}) = \widehat{w}(\mathbf{k}) + \sum_{\boldsymbol{\ell} \neq 0} \widehat{w}(\mathbf{k} + \boldsymbol{\ell}(2N+1)), \qquad |\mathbf{k}| \leq N,$$

<sup>&</sup>lt;sup>1</sup>There is a slight difference between the formulae based on an even and an odd number of points; we chose to continue with the slightly simpler notations associated with an odd number of points.

where summation runs over all *d*-tuples,  $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_d) \neq 0$ . It shows that all the Fourier coefficients with wavenumber  $\mathbf{k}[mod(2N+1)]$  are "aliased" into the same discrete Fourier coefficient,  $\tilde{w}_{\mathbf{k}}$ . It follows that the interpolation error consists of two main contributions,

$$(I - \psi_N)[w] \equiv (I - P_N)[w] + A_N[w]$$

where in addition to the truncation error  $(I - P_N)[w]$  in (1.1), we now have the *aliasing* error,

(1.3) 
$$A_N[w](\mathbf{x}) := \sum_{|\mathbf{k}| \le N} \Big( \sum_{|\boldsymbol{\ell}| \ge 1} \widehat{w}(\mathbf{k} + \boldsymbol{\ell}(2N+1)) \Big) e^{i\mathbf{k}\cdot\mathbf{x}}.$$

Both,  $(I - P_N)[w]$  and  $A_N[w]$ , involve high modes,  $\widehat{w}(\mathbf{p})$ ,  $|\mathbf{p}| \ge N$ . Consequently, if the function  $w(\cdot)$  is sufficiently smooth then they have exactly the same spectrally small size, e.g. [Tad94, §2.2]

$$||A_N[w]||_{H^s} \lesssim ||(I - P_N)[w]||_{H^s} \lesssim N^{s-r} ||w||_{H^r}, \qquad r > s > \frac{d}{2}.$$

Since the truncation error is orthogonal to the computational N-space whereas the aliasing error is not the situation is different if w lacks smoothness. Then aliasing and truncation errors may have a completely different influence on the question of computational stability. One such case is encountered with the stability question of spectral vs. pseudo-spectral approximations of hyperbolic equations.

# 2. Fourier method for linear equations — weak instability for $L^2$ -data

2.1. The spectral approximation: stability and convergence. We want to solve the  $2\pi$ -periodic scalar hyperbolic equation

(2.1) 
$$\frac{\partial}{\partial t}u(x,t) + \frac{\partial}{\partial x}(q(x)u(x,t)) = 0, \qquad x \in \mathbb{T}([0,2\pi)), \quad q \in C^1[0,2\pi],$$

subject to prescribed initial conditions,  $u(\cdot, 0)$ , by the spectral Fourier method. To this end we approximate the spectral projection of the exact solution,  $P_N u(\cdot, t)$ , using an *N*degree polynomial,  $u_N(x,t) = \sum_{|k| \leq N} \hat{u}_k(t) e^{ikx}$ , which is governed by the semi-discrete approximation, [Or72, KO72, GO77]

(2.2) 
$$\frac{\partial}{\partial t}u_N(x,t) + \frac{\partial}{\partial x}P_N[q(x)u_N(x,t)] = 0.$$

The approximation is realized as a convolution in Fourier space

$$\frac{d}{dt}\widehat{u}_k(t) = ik \sum_{|j| \le N} \widehat{q}(k-j)\widehat{u}_j(t), \qquad k = -N, \dots, N,$$

which amounts to a system of (2N+1) ODEs for the computed  $\widehat{\mathbf{u}}(t) := (\widehat{u}_{-N}(t), \dots, \widehat{u}_{N}(t))^{\top}$ .

The  $L^2$ -stability of (2.2) is straightforward: though the truncation error which enters (2.2),  $\partial_x (I - P_N)[q(x)u_N(x)]$  need not be small, it is orthogonal to the N-space, and hence,

(2.3) 
$$\frac{1}{2}\frac{d}{dt}\|u_N(\cdot,t)\|_{L^2}^2 = -\int_0^{2\pi} u_N \frac{\partial}{\partial x} P_N[q(x)u_N] dx = \int \frac{\partial u_N}{\partial x} P_N[q(x)u_N] dx$$
$$= \int \frac{\partial u_N}{\partial x} q(x)u_N dx = -\frac{1}{2} \int q'(x)u_N^2 dx$$
$$\leq \frac{1}{2} \max_x |q'(x)| \times \|u_N(\cdot,t)\|_{L^2}^2.$$

This yields the  $L^2$ -stability bound,

(2.4) 
$$\|u_N(\cdot,t)\|_{L^2}^2 \le e^{q'_{\infty}t} \|u_N(\cdot,0)\|_{L^2}^2, \qquad q'_{\infty} := \max_x |q'(x)|.$$

To convert this stability bound into a spectral convergence rate estimate, consider the difference between the spectral method (2.2) and the  $P_N$  projection of the underlying equation (2.1): one finds that  $e_N := u_N - P_N u$ , satisfies the error equation

$$\frac{\partial}{\partial t}e_N(x,t) + \frac{\partial}{\partial x}P_N[q(x)e_N(x,t)] = -\frac{\partial}{\partial x}P_N[q(x)(I-P_N)[u](x,t)].$$

The  $L^2$ -stability (2.4) implies the error estimate,

$$\int |u_N(x,t) - P_N u(x,t)|^2 dx \lesssim e^{q'_\infty t} \left( \|(I - P_N)u(\cdot,0)\|_{L^2}^2 + N^2 \max_{\tau \leq t} \|(I - P_N)[u](\cdot,\tau)\|_{L^2}^2 \right) dx$$

This quantifies the *spectral convergence* of the Fourier method (2.2): the convergence rate increases together with the increasing order of smoothness of the solution,

(2.5) 
$$\|u_N(\cdot,t) - u(\cdot,t)\|_{L^2} \lesssim e^{\frac{1}{2}q'_{\infty}t} \left( N^{-s} \|u(\cdot,0)\|_{H^s} + N^{1-s} \max_{\tau \le t} \|u(\cdot,\tau)\|_{H^s} \right), \quad s > 1.$$

In practice, one recovers exponential convergence for analytic solutions (and root-exponential convergence for more general Gevrey data).

2.2. The pseudo-spectral approximation: aliasing and weak stability. We now consider pseudo-spectral Fourier approximation of (2.1). As before, we use an N-degree polynomial,  $u_N(x,t) = \sum_{|k| \le N} \hat{u}_k(t) e^{ikx}$ , as an approximation for  $\psi_N u(\cdot, t)$ , which is governed by the semi-discrete approximation, [KO72, GO77],

(2.6) 
$$\frac{\partial}{\partial t}u_N(x,t) + \frac{\partial}{\partial x}\psi_N[q(x)u_N(x,t)] = 0$$

This equation can be realized in physical space

$$\frac{d}{dt}u_N(x_j,t) = \sum_{k=-N}^{N} ik \widetilde{(qu_N)}_k e^{ikx_j}, \quad \widetilde{(qu_N)}_k = \frac{h}{2\pi} \sum_{\nu=0}^{2N} q(x_\nu) u_N(x_\nu) e^{-ikx_\nu}.$$

It amounts to a system of (2N+1) ODEs for the computed gridvalues  $\mathbf{u}(t) := (u(x_0, t), \dots, u(x_{2N}, t))^\top$ 

$$\frac{d}{dt}\mathbf{u}(t) = DQ\mathbf{u}(t), \qquad D_{jk} = \frac{(-1)^{j-k}}{2\sin\left(\frac{x_j - x_k}{2}\right)}, \quad Q = \begin{pmatrix} q(x_0) & 0 & \dots & 0\\ 0 & \ddots & 0 & \vdots\\ \vdots & & \ddots & 0\\ 0 & 0 & \dots & q(x_{2N}) \end{pmatrix}$$

Here D is the Fourier differentiation matrix and Q signifies pointwise multiplication with q(x).

To examine the stability of (2.6) we repeat the usual  $L^2$ -energy argument for the spectral approximation in (2.2): decompose  $\psi_N = P_N + A_N$ , to find

(2.7) 
$$\frac{1}{2}\frac{d}{dt}\|u_N(\cdot,t)\|_{L^2}^2 = \underbrace{\int u_N \frac{\partial}{\partial x} P_N[q(x)u_N]dx}_{\text{a bounded term-(2.3)}} + \underbrace{\int u_N \frac{\partial}{\partial x} A_N[q(x)u_N]dx}_{\text{contribution of aliasing}}$$

The first term on the right consists of truncation error which, by (2.3), does not exceed  $\leq ||u_N(\cdot, t)||^2$ . Thus, the stability of the Fourier approximation (2.6) depends solely on the

aliasing contributions,  $A_N[q(x)u_N]$ : using (1.3) to expand the second term on the right, we find

(2.8) 
$$\int u_N \frac{\partial}{\partial x} A_N[q(x)u_N] dx = 2\pi i \sum_{|j|,|k| \le N} \overline{\widehat{u}}_j(t) \widehat{u}_k(t) (j-k) \cdot \sum_{\ell \ne 0} \widehat{q} (j-k+\ell(2N+1)).$$

Observe that the terms on the right,  $\sum_{\ell \neq 0} \hat{q} (j - k + \ell(2N + 1))$ , are of order  $\mathcal{O}(N)$  for  $|j - k| \sim 2N, \ell = \pm 1$ , and this can occur only for high wavenumbers,  $|j| \sim |k| \sim N$ . Thus, there is possible  $\mathcal{O}(N)$  amplification of the high Fourier modes,  $|\hat{u}_j(t)|, |j| \sim N$ . Unfortunately, these Fourier modes need not be small due to lack of apriori smoothness, and aliasing may render the Fourier method as unstable.

Indeed, when q(x) changes sign, the exact solution of (2.1) develops large gradients and consequently, the Fourier method does experience spurious oscillations precisely because of aliasing errors which are ignited due to lack of smoothness. To demonstrate the exact mechanism of this type of instability<sup>2</sup>, we consider the example  $q(x) = \sin(x)$ ,

(2.9) 
$$\frac{\partial}{\partial t}u_N(x,t) + \frac{\partial}{\partial x}\psi_N[\sin(x)u_N(x,t)] = 0$$

The analysis follows [GHT94]. Fourier transform of (2.9) yields

$$\frac{d}{dt}\widehat{u}_k(t) = \frac{k}{2}[\widehat{u}_{k-1}(t) - \widehat{u}_{k+1}(t)], \quad k = -N, \dots, N$$

and its imaginary part,  $b_k(t) := \Im \widetilde{u}_k(t)$ , reads

(2.10) 
$$\frac{d}{dt}b_k(t) = \frac{k}{2}\left[b_{k-1}(t) - b_{k+1}(t)\right], \quad b_{N+1} + b_N = 0.$$

The last set of ODEs is at the heart of matter. A straightforward energy estimate yields the *lower-bound*  $\frac{d}{dt} \|\mathbf{b}(t)\|^2 \ge -\|\mathbf{b}(t)\|^2 + Nb_N^2(t)$  for  $\mathbf{b}(t) := (b_1(t), \dots, b_N(t))^{\top}$ . Does  $b_N(t)$ grow? on the one hand, numerical simulations in Figure 2.1 show that it does. On the other hand, *if* the solution is smooth enough, then aliasing errors are negligible: in Figure 2.2 for example, the solution subject to initial data  $|\hat{u}_k(0)| \sim |k|^{-3}$  remain smooth and the spurious mode,  $|b_N(t)| \sim N^{-2}$ , decay sufficiently fast so that  $\mathbf{b}(t)$  remains bounded.



FIGURE 2.1.  $b_k(t) = \Im \hat{u}_k(t)$ : Fourier method subject to  $\hat{u}_k(0) = x_k^3(\pi - x_k)^3/20$ . (a) (t, N) = (0.1, 200); (b) (t, N) = (1., 100); (c) (t, N) = (1., 200).

<sup>&</sup>lt;sup>2</sup>To demonstrate the instability of the Fourier method (2.6), one needs to consider q(x) which changes sign. Otherwise, if q(x) is, say, positive, then DQ is similar to the anti-symmetric matrix  $\mathcal{A} := \sqrt{Q}D\sqrt{Q}$ and stability follows since  $exp(\mathcal{A}t)$  is unitary in  $\mathbb{C}^{2N+1}$  for the scalar product  $(QU_N, V_N)$  [KO72].



FIGURE 2.2.  $\Im \hat{u}_k(t)$  computed at t = 3: Fourier solution of  $u_t = (\sin(x)u)_x$ ,  $\hat{u}_k(0) \sim k^{-3}$ . (a) with N = 100; (b) with N = 200; (c) with N = 800:  $b_N(t) \sim \pm N^{-2}$ .

In practice, however, even if the solution of (2.1) remains smooth, it develops large gradients of order  $|\hat{\mathbf{u}}(t)| \sim exp(q'_{\infty}t)$  when q(x) changes sign. These large gradients require  $N \gg e^t$  modes in order to resolve these gradients; otherwise, the exact solution  $u(\cdot, t)$ remains under-resolved by the pseudo-spectral Fourier approximation. Observe that the Fourier method requires an increasingly large number of modes before it can resolve the underlying solution,  $u(\cdot, t)$ . Without it, the under-resolved Fourier approximation contains  $\mathcal{O}(1)$  high modes,  $|\hat{u}_N(t)| \sim |b_N(t)|$ , amplified by a factor of order  $\mathcal{O}(N)$ , yielding the spurious oscillations noticeable in Figure 2.3. Thus, aliasing errors cause the Fourier solution to grow due to *lack of resolution*. The precise growth is "encoded" in the Fourier equations whose imaginary part is governed by (2.10). To this end, set  $v_k(t) := (-1)^k b_k(t)$ . Then (2.10) is converted into

$$\frac{d}{dt}v_k(t) = x_k \frac{v_{k+1}(t) - v_{k-1}(t)}{2\Delta x}, \quad v_{N+1}(t) = v_N(t), \quad x_k := k\Delta x \in [0, 1], \ \Delta x := h/\pi.$$

This can be viewed as an approximation to the linear equation  $\partial_t v(x,t) = x \partial_x v(x,t)$ ,  $0 \le x \le 1$  augmented with the boundary condition  $\partial_x v(1,t) = 0$ . This is an *ill-posed* problem due to the extrapolation at the inflow boundary x = 1, and consequently, its numerical approximation grows  $||b_N(t)|| = ||v_N(t)|| \sim \sqrt{N}$ , e.g., [Tad83]. The detailed analysis carried out in [GHT94] shows that there is a *weak* instability, where  $\sim 1 - e^{-t}$  fraction of the highest modes experience amplification of order  $\mathcal{O}(N)$ , which ends with the stability estimate

$$||u_N(\cdot,t)||_{L^2} \lesssim e^{Cq'_{\infty}t} N ||u_N(\cdot,0)||_{L^2}.$$

The corresponding error estimate for the pseudo-spectral approximation reads [GHT94, theorem 4.1]

$$\|u_N(\cdot,t) - \psi_N u(\cdot,t)\|_{L^2} \lesssim e^{C_s q'_{\infty} t} \left( N^{1-s} \|u(\cdot,0)\|_{H^s} + N^{2-s} \max_{\tau \le t} \|u(\cdot,\tau)\|_{H^s} \right), \qquad s > 2,$$

reflecting the loss of power on N when compared with the spectral estimate (2.5).

2.3. De-aliasing: the 2/3 smoothing method and strong stability. One way to regain the stability of the pseudo-spectral Fourier method in (2.9) is to set  $\hat{u}_N(t) \equiv 0$  which prevents the growth of  $b_N(t) \equiv 0$ . Thus, removing the last mode stabilize the pseudospectral method in the special case of  $q(x) = \sin(x)$ . The hyperbolic equation (2.1) with a general



FIGURE 2.3.  $\Im \hat{u}_k(t=0.4)$  of  $u_N(\cdot, t)$  with N = 64, N = 128 and N = 512 modes.

q(x) follows along the lines of [Tad87]. We return to the aliasing term (2.8),

(2.11) 
$$\int u_N(x,t) \frac{\partial}{\partial x} A_N[q(\cdot)u_N](x,t) dx$$
$$= 2\pi i \sum_{|j|,|k| \le N} \overline{\widehat{u}}_j(t) \widehat{u}_k(t) (j-k) \cdot \sum_{\ell \ne 0} \widehat{q} (j-k+\ell(2N+1))$$

As noted before, the  $\mathcal{O}(N)$  growth of the high Fourier modes,  $|\hat{u}_j(t)|$ ,  $|j| \sim N$ , need not be small due to lack of apriori smoothness. We can circumvent this difficulty if we remove these modes by setting a fixed portion of the spectrum to be zero. For example, assume that we truncate the last 1/3 of the modes of  $u_N$  (any other fixed fraction of N will do). To this end, we use a *smoothing operator* S which is activated only on the first  $\frac{2}{3}N$  modes while removing the top 1/3N of the modes. We up with the so-called 2/3 pseudo-spectral Fourier method,

(2.12a) 
$$\frac{\partial}{\partial t}u_N(x,t) + \frac{\partial}{\partial x}\psi_N[q(\cdot)\mathcal{S}u_N](x,t) = 0, \qquad \mathcal{S}u_N := \sum_{|k| \le \frac{2}{3}N} \sigma_k \widehat{u}_k(t)e^{ikx};$$

To retain spectral accuracy, the smoothing factors  $\sigma_k \in (0, 1]$  do not change a fixed portion of the lower spectrum

(2.12b) 
$$\sigma_k \begin{cases} \equiv 1, & |k| \le \frac{1}{3}N \\ \in (0,1], & \frac{1}{3}N < |k| < \frac{2}{3}N \end{cases}$$

The  $L^2$ -stability of the 2/3 method follows along the lines of the spectral stability in (2.3); integrating (2.12a) against  $Su_N$  we find

(2.13a) 
$$\left| \int (\mathcal{S}u_N) \frac{\partial}{\partial x} P_N[q(\cdot)\mathcal{S}u_N] dx \right| \le \frac{1}{2} q'_{\infty} \|\mathcal{S}u_N(\cdot,t)\|_{L^2}^2, \qquad q'_{\infty} = \max_x |q'(x)|.$$

The aliasing contribution in the 2/3 method is bounded (and in fact negligible for  $q \in C^s, s > 1$ ): following (2.11)

$$\int (\mathcal{S}u_N) \frac{\partial}{\partial x} A_N[q(\cdot)(\mathcal{S}u_N)] dx = 2\pi i \sum_{|j| \le \frac{2}{3}N} \sum_{|k| \le \frac{2}{3}N} \sigma_k \sigma_j \overline{\widehat{u}}_j(t) \widehat{u}_k(t) (j-k) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) \ge \frac{2}{3}N}_{\ell \ne 0} (j-k+\ell(2N+1)) \cdot \sum_{\ell \ne 0} \widehat{q} \underbrace{(j-k+\ell(2N+1)) + 2}\widehat{q} \underbrace{(j-k+\ell(2N+1)) + 2$$

observe that the terms involved in the inner summation on the right are now restricted to high wavenumbers,  $|j - k + \ell(2N + 1)| \ge \frac{2}{3}N$  so that  $|\hat{q}(j - k + \ell(2N + 1))| \le ||q||_{C^r}N^{-r}$ . Hence

(2.13b) 
$$\left| \int (\mathcal{S}u_N) \frac{\partial}{\partial x} A_N[q(\cdot)(\mathcal{S}u_N)] dx \right| \lesssim \|q\|_{C^r} N^{1-r} \times \|\mathcal{S}u_N\|^2, \qquad r \ge 1$$

Combining the last two inequalities (2.13a) and (2.13b) with r = 1, we arrive at

$$\frac{1}{2}\frac{d}{dt}\int (\mathcal{S}u_N)(x,t)u_N(x,t)dx = -\int (\mathcal{S}u_N)\frac{\partial}{\partial x}(P_N + A_N)[q(\cdot)(\mathcal{S}u_N)]dx \le Cq'_{\infty}\|\mathcal{S}u_N(\cdot,t)\|_{L^2}^2.$$

Thus, by activating the smoothing operator we removed aliasing errors and the resulting 2/3 de-aliased pseudo-spectral method (2.12) regained the *weighted*  $L^2$ -stability

$$\|u_N(\cdot,t)\|_{L^2_{\mathcal{S}}}^2 \le e^{2Cq'_{\infty}t} \|u_N(\cdot,0)\|_{L^2_{\mathcal{S}}}^2, \qquad \|w(\cdot,t)\|_{L^2_{\mathcal{S}}}^2 := \int (\mathcal{S}w)(x,t)w(x,t)dx = 2\pi \sum_{|k| \le \frac{2}{3}N} \sigma_k |\widehat{w}_k(t)|^2.$$

The corresponding error equation for  $e_N := Su_N - Su$  reads (we skip the details)

$$\frac{\partial}{\partial t}e_N + \frac{\partial}{\partial x} \left( \mathcal{S}[qe_N] \right) = -\frac{\partial}{\partial x} \mathcal{S}[q(x)(I - \mathcal{S})[u](x, t)],$$

and the spectral convergence rate, (2.5), follows: for s > 1 there exists a constant,  $C = C_s$  such that

$$\|u_N - u\|_{L^2_{\mathcal{S}}} \lesssim e^{C_s q'_{\infty} t} \left( N^{-s} \|u(\cdot, 0)\|_{H^s} + N^{1-s} \max_{\tau \le t} \|u(\cdot, \tau)\|_{H^s} \right), \quad s > 1.$$

2.4. Spectral accuracy and propagation of discontinuities. Hyperbolic equations propagates  $H^s$  regularity:  $||u(\cdot,t)||_{H^s} \leq e^{C_s t} ||u(\cdot,0)||_{H^s} < \infty$ . Thus, the convergence statement in (2.5) implies spectral convergence of the spectral Fourier method and 2/3 Fourier method for  $H^s$ -smooth initial data. However, when the initial data is *piecewise smooth*, the exact solution propagates discontinuities along characteristics, and the (pseudo-)spectral approximations of jump discontinuities in  $u(\cdot, t)$  produces spurious Gibbs oscillations, [Tad07]. Nevertheless, thanks to the  $H^s$ -stability of the spectral Fourier method and the 2/3 pseudospectral methods,  $||u_N(\cdot,t)||_{\dot{H}^s} \leq e^{C_s |q| \propto t} ||u_N(\cdot,0)||_{\dot{H}^s}$ , measured in the *weak* topology of s < 0, the (pseudo-)spectral approximations still propagate accurate information of the smooth portions of the exact solution. This is realized in terms of the convergence rate (we skip the details)

$$\|u_N - u\|_{H^r} \lesssim e^{C_s q'_{\infty} t} \left( N^{r-s} \|u(\cdot, 0)\|_{H^s} + N^{1+r-s} \max_{\tau \le t} \|u(\cdot, \tau)\|_{H^s} \right), \qquad r < s - 1 < -1.$$

It follows that one can pre- and post-process  $u_N(\cdot, t)$  to recover the *pointvalues* of  $u(\cdot, t)$  within spectral accuracy, away from the singular set of the solution, [MMO78, ML78, AGT86]. The point to note here is that even the Fourier projections of the exact solution,  $P_N u(\cdot, t)$  and  $\psi_N u(\cdot, t)$  are at most first-order accurate due to Gibbs oscillations; the post-processing of the computed  $u_N$  is realized by its smoothing using a proper  $\sigma$ -mollifier

(2.12b) (or see (3.6c) below), which does both — retains the stability and recovers the spectrally accurate resolution content of the Fourier method.

## 3. Fourier method for Burgers equation: convergence for smooth solutions

We now turn our attention to spectral and pseudo-spectral approximations of nonlinear problems. Their spectral accuracy often make them the method of choice for simulations where the highest resolution is sought for a given number of degrees of freedom. We begin with the prototypical example for quadratic nonlinearities, the inviscid Bugers' equation,

(3.1) 
$$\frac{\partial}{\partial t}u(x,t) + \frac{1}{2}\frac{\partial}{\partial x}u^2(x,t) = 0, \qquad x \in \mathbb{T}([0,2\pi)),$$

subject to  $2\pi$ -periodic boundary conditions and prescribed initial conditions, u(x, 0). In this section we show that as long as the solution of Burgers equation remains smooth for a time interval  $t \leq T_c$ , the spectral and 2/3 de-aliased pseudo-spectral approximations converge to the exact solution with spectral accuracy.

3.1. The spectral approximation of Burgers equation. The spectral approximation of (3.1),  $u_N(x,t) = \sum \hat{u}_k(t)e^{ikx}$ , is governed by,

(3.2) 
$$\frac{\partial}{\partial t}u_N(x,t) + \frac{1}{2}\frac{\partial}{\partial x}\left(P_N\left[u_N^2\right](x,t)\right) = 0, \qquad 0 \le x \le 2\pi.$$

The evaluation of the quadratic term on the right is carried out using convolution and (3.2) amounts to a nonlinear system of (2N+1) ODEs for  $\widehat{\mathbf{u}}(t) = (\widehat{u}_{-N}(t), \dots, \widehat{u}_{N}(t))^{\top}$ .

**Theorem 3.1 (Spectral convergence for smooth solutions of Burgers' equations).** Assume that for  $0 < t \le T_c$ , the solution of the Burgers equation (2.1) is smooth,  $u(\cdot, t) \in L^{\infty}([0, T_c], C^{1+\alpha}(0, 2\pi])$ . Then, the spectral method (3.2) converges in  $L^{\infty}([0, T_c], L^2(0, 2\pi])$ ,

$$||u_N(\cdot, t) - u(\cdot, t)||_{L^2} \to 0, \qquad 0 \le t \le T_c$$

Moreover, the following spectral convergence rate estimate holds for all  $s > \frac{3}{2}$ ,

$$\|u_N(\cdot,t) - u(\cdot,t)\|_{L^2}^2 \lesssim e^{\int_0^t |u_x(\cdot,\tau)|_{\infty} d\tau} \left( N^{-2s} \|u(\cdot,0)\|_{H^s}^2 + N^{\frac{3}{2}-s} \max_{\tau \le t} \|u(\cdot,\tau)\|_{H^s} \right), \ s > \frac{3}{2}.$$

*Proof.* We rewrite the spectral approximation (3.1) in the form

$$\frac{\partial}{\partial t}u_N + \frac{\partial}{\partial x}\frac{u_N^2}{2} = \frac{1}{2}\frac{\partial}{\partial x}(I - P_N)[u_N^2]$$

The corresponding energy equation reads

(3.3) 
$$\frac{\partial}{\partial t}\frac{u_N^2}{2} + \frac{\partial}{\partial x}\frac{u_N^3}{6} = \frac{u_N}{2}\frac{\partial}{\partial x}(I - P_N)[u_N^2].$$

Integration yields the energy balance

$$\frac{1}{2}\frac{d}{dt}\int u_N^2(x,t)dx = \frac{1}{2}\int u_N\partial_x(I-P_N)[u_N^2]dx =: \mathcal{I}_1$$

The term on the right vanishes by orthogonality,  $\mathcal{I}_1 = -\frac{1}{2} \int \frac{\partial u_N}{\partial x} (I - P_N) [u_N^2] dx = 0$ , and hence the solution is  $L^2$ -conservative,

(3.4) 
$$\|u_N(\cdot,t)\|_{L^2} = \|u_N(\cdot,0)\|_{L^2}.$$

Next, we integrate  $(u_N - u)^2 \equiv |u_N|^2 - |u|^2 - 2u(u_N - u)$ : after discarding all terms which are in divergence form, we are left with

$$\frac{1}{2}\frac{d}{dt}\int (u_N - u)^2 dx = \frac{d}{dt}\int \left(\frac{|u_N|^2}{2} - \frac{|u|^2}{2} - u(u_N - u)\right) dx$$
$$= \frac{1}{2}\int u_N \partial_x (I - P_N)[u_N^2] dx - \int \partial_t (u(u_N - u)) dx =: \mathcal{I}_1 + \mathcal{I}_2.$$

Recall that  $\mathcal{I}_1$  vanishes. As for the second term  $\mathcal{I}_2$ , we decompose it into two terms,

$$\mathcal{I}_2 = \int \partial_t \big( u(u_N - u) \big) dx \equiv \int u_t(u_N - u) dx + \int u(\partial_t u_N - \partial_t u) dx,$$

and using (3.1), (3.2) and (3.3) to convert time derivatives to spatial ones, we find

$$\begin{aligned} \mathcal{I}_{2} &= -\int u u_{x}(u_{N}-u) dx - \int u \partial_{x} \left(\frac{u_{N}^{2}}{2} - \frac{u^{2}}{2}\right) dx + \frac{1}{2} \int u \partial_{x}(I-P_{N})[u_{N}^{2}] dx \\ &= -\int u u_{x}(u_{N}-u) dx + \int u_{x} \left(\frac{u_{N}^{2}}{2} - \frac{u^{2}}{2}\right) dx - \frac{1}{2} \int u_{x}(I-P_{N})[u_{N}^{2}] dx \\ &= \int u_{x} \left(\frac{u_{N}^{2}}{2} - \frac{u^{2}}{2} - u(u_{N}-u)\right) dx - \frac{1}{2} \int u_{x}(I-P_{N})[u_{N}^{2}] dx. \end{aligned}$$

Eventually, we end up with

(3.5a) 
$$\frac{1}{2}\frac{d}{dt}\int |u_N(x,t) - u(x,t)|^2 dx \le \frac{|u_x(\cdot,t)|_{L^{\infty}}}{2}\int |u_N(x,t) - u(x,t)|^2 dx - \frac{1}{2}e_N,$$

where the error term,  $e_N$ , is given by

(3.5b) 
$$e_N := \int u_N^2 (I - P_N) [u_x] dx$$

Observe that under the hypothesis  $u_x \in L^{\infty}_t C^{0,\alpha}_x$ , and hence by Jackson's bound [DL93] and the  $L^2$ -bound (3.4) one has

$$|e_N(t)| \lesssim \max_x |(I - P_N)[u_x(x, t)]| \cdot ||u_N||_{L^2}^2 \lesssim \frac{\ln N}{N^{\alpha}} ||u_N(\cdot, 0)||_{L^2}^2 \to 0.$$

With (3.5) one obtains,

$$\int |u_N(x,t) - u(x,t)|^2 dx \leq e^{U'_{\infty}(t;0)} \int |u_N(x,0) - u(x,0)|^2 dx + \int_0^t e^{U'_{\infty}(t;\tau)} |e_N(\tau)| d\tau, \qquad U'_{\infty}(t;\tau) := \int_{s=\tau}^t |u_x(\cdot,s)|_{\infty} ds.$$

and convergence follows. Moreover, with  $u_N(\cdot, 0) = P_N u(\cdot, 0)$  we end up with spectral convergence rate estimate

$$\begin{split} \int |u_N(x,t) - u(x,t)|^2 dx \\ \lesssim e^{\int_0^t |u_x(\cdot,\tau)|_\infty d\tau} \left( N^{-2s} \|u(\cdot,0)\|_{H^s}^2 + N^{\frac{3}{2}-s} \max_{\tau \le t} \|u(\cdot,\tau)\|_{H^s} \right), \qquad s > \frac{3}{2}. \end{split}$$

12

3.2. The 2/3 de-aliasing pseudo-spectral approximation of Burgers equation. Convolutions can be avoided using the pseudo-spectral Fourier method, [KO72, GO77],

$$\frac{\partial}{\partial t}u_N(x,t) + \frac{1}{2}\frac{\partial}{\partial x}\left(\psi_N\left[u_N^2\right](x,t)\right) = 0, \qquad x \in \mathbb{T}([0,2\pi)).$$

Observe that (3.1) is satisfied *exactly* at the gridpoints  $x_{\nu}$ ,

$$\frac{d}{dt}u_N(x_\nu,t) + \frac{1}{2}\frac{\partial}{\partial x}\left(\psi_N\left[u_N^2\right](x,t)\right)\Big|_{x=x_\nu} = 0, \qquad \nu = 0, 1, \dots, 2N.$$

The resulting system of (2N+1) nonlinear equations for  $\mathbf{u}(t) = (u(x_0, t), \dots, u(x_{2N}, t))^{\top}$  can be then integrated in time by standard ODE solvers. The pseudo-spectral approximation introduces aliasing errors. To eliminate these errors, we consider the 2/3 de-aliasing Fourier method, consult (2.12a)

(3.6a) 
$$\frac{\partial}{\partial t}u_N(x,t) + \frac{1}{2}\frac{\partial}{\partial x}\left(\psi_N\left[(\mathcal{S}u_N)^2\right](x,t)\right) = 0, \quad x \in \mathbb{T}([0,2\pi)),$$

where  $Su_N$  denotes a *smoothing* operator of the form

(3.6b) 
$$Su_N := \sum_{|k| \le \frac{2}{3}N} \sigma_k \widehat{u}_k(t) e^{ikx}, \qquad \widehat{u}_k(t) = \frac{h}{2\pi} \sum_{\nu=0}^{2N} u_N(x_\nu, t) e^{-ikx_\nu}.$$

The smoothing operator S is dictated by the smoothing factors,  $\{\sigma_k\}_{|k| \leq \frac{2}{3}N}$ , which truncates modes with wavenumbers  $|k| > \frac{2}{3}N$  while leaving a fixed portion — say, the first 1/3 of the spectrum, viscous-free. This is the same smoothing operator  $Su_N$  we considered already in the *linear* 2/3 method (2.12a). In typical cases, one may employ a smoothing mollifier,  $\sigma(\cdot) \in C^{\infty}(0, 1)$ , setting

(3.6c) 
$$\sigma_{k} = \sigma\left(\frac{|k|}{N}\right), \quad \sigma(\xi) \begin{cases} \equiv 1, & \xi \leq \frac{1}{3}, \\ \in (0,1), & \frac{1}{3} < \xi < \frac{2}{3}, \\ \equiv 0, & \frac{2}{3} \leq \xi \leq 1. \end{cases}$$

This is the 2/3 de-aliasing Fourier method which is often advocated for spectral computations, in particular those involving quadratic nonlinearities, [HL07, OHFS10, Kerr93, Kerr05].

In what sense does the 2/3 method remove aliasing errors? to make precise the de-aliasing aspect of (3.6), consider the 2/3 truncated solution  $u_m := Su_N$ . Here we emphasize that we are dealing with the smoothed solution,  $u_m$ , of degree  $m := \frac{2}{3}N$ . Observing that truncation commute with differentiation, we find

(3.7) 
$$\frac{\partial}{\partial t}u_m(x,t) + \frac{1}{2}\frac{\partial}{\partial x}\mathcal{S}\left(\psi_N[u_m^2]\right)(x,t) = 0, \qquad deg(u_m) = m := \frac{2}{3}N.$$

We now come to the key point behind the removal of aliasing in quadratic nonlinearities: since  $\hat{u}_m(k) = 0$  for  $|k| > \frac{2}{3}N$  then  $\widehat{u_m^2}(k) = 0$  for  $|k| > \frac{4}{3}N$  hence  $\widehat{u_m^2}(k + \ell(2N + 1)) = 0$ for  $|k| \le \frac{2}{3}N, \ell \ne 0$ ; consequently, since the smoothing operator S acts only on the first  $\frac{2}{3}N$ mode,  $S(A_N u_m^2) \equiv 0$ , and we conclude

$$\mathcal{S}(\psi_N[u_m^2])(x,\cdot) \equiv \mathcal{S}((P_N + A_N)[u_m^2])(x,\cdot) = \mathcal{S}(P_N[u_m^2])(x,\cdot) \equiv \mathcal{S}u_m^2(x,\cdot).$$

We summarize by stating the following.

**Corollary 3.2.** Consider the 2/3 de-aliasing Fourier method (3.6) then its 2/3 smoothed solution,  $u_m := Su_N$ , satisfies

(3.8) 
$$\frac{\partial}{\partial t}u_m(x,t) + \frac{1}{2}\frac{\partial}{\partial x}\mathcal{S}[u_m^2](x,t) = 0, \qquad \mathcal{S}w = \sum_{|k| \le m} \sigma_k \widehat{w}_k e^{ikx}, \ m = \frac{2}{3}N_k$$

Thus, by truncating the top 1/3 of the modes, we de-aliased the Fourier method, (3.6a), in the sense that (3.8) does *not* involve any aliasing errors: only truncation errors,  $(I - S)[u_m^2]$ are involved. Indeed, the formulation of 2/3 method in (3.8) resembles the *m*-mode spectral method (3.2). The only difference is due to the fact that unless  $\sigma_k \equiv 1$ , the smoothing operator S is not a projection<sup>3</sup>

The following theorem shows that as long as the Burgers solution remains smooth, the 2/3 de-aliasing Fourier method is stable and enjoys spectral convergence.

**Theorem 3.3 (Spectral convergence of the** 2/3 **method for smooth solutions).** Assume that for  $0 < t \leq T_c$ , the solution of the Burgers equation (2.1) is smooth,  $u(\cdot, t) \in L^{\infty}([0, T_c], C^{1+\alpha}(0, 2\pi])$ . Then, the 2/3 de-aliasing method (3.6) converges in  $L^{\infty}([0, T_c], L^2(0, 2\pi])$ ,

$$||u_m(\cdot, t) - u(\cdot, t)||_{L^2} \to 0, \qquad 0 \le t \le T_c,$$

and the following spectral convergence rate estimate holds

$$\|u_N(\cdot,t) - u(\cdot,t)\|_{L^2}^2 \lesssim e^{\int_0^t |u_x(\cdot,\tau)|_\infty d\tau} \left( N^{-2s} \|u(\cdot,0)\|_{H^s}^2 + N^{\frac{3}{2}-s} \max_{\tau \le t} \|u(\cdot,\tau)\|_{H^s} \right), \ s > \frac{3}{2}$$

*Proof.* We start with (3.8)

$$\frac{\partial}{\partial t}u_m(x,t) + \frac{1}{2}\frac{\partial}{\partial x}\Big(\mathcal{S}[u_m^2](x,t)\Big) = 0.$$

Since S need not be a projection, there is no  $L^2$ -energy conservation for the 2/3 smoothed solution  $u_m$ . Instead, we integrate against  $u_N$  to find that the corresponding energy balance reads

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\int u_N(x,t)u_m(x,t)dx &= -\frac{1}{2}\int u_N\frac{\partial}{\partial x}\mathcal{S}[u_m^2](x,t)dx\\ &= -\frac{1}{2}\int\frac{\partial}{\partial x}(\mathcal{S}u_N)u_m^2(x,t)dx = \frac{1}{6}\int\frac{\partial}{\partial x}u_m^3dx = 0, \end{aligned}$$

and hence the solution conserve the weighted  $L_{\mathcal{S}}^2$ -norm, (3.9)

$$\|u_m(\cdot,t)\|_{L^2_{\mathcal{S}}}^2 = \|u_N(\cdot,0)\|_{L^2_{\mathcal{S}}}^2, \qquad \|u_N(\cdot,t)\|_{L^2_{\mathcal{S}}}^2 := \int (\mathcal{S}u_N)u_N dx = 2\pi \sum_{|k| \le \frac{2}{3}N} \sigma_k |\widehat{u}_k(t)|^2.$$

We proceed along the lines of the spectral proof in theorem 3.1, integrating  $|u_m - u|^2 \equiv |u_m|^2 - |u|^2 - 2u(u_m - u)$ : after discarding all terms which are in divergence form, we are

$$\frac{\partial}{\partial t}u_m(x,t) + \frac{1}{2}\frac{\partial}{\partial x}P_m[u_m^2](x,t) = 0, \qquad |k| \le \frac{2}{3}N.$$

<sup>&</sup>lt;sup>3</sup>When  $\sigma_k \equiv 1$ , then  $S = P_{\frac{2}{3}N}$  and the 2/3 method coincides with the spectral Fourier method (3.2) with  $m = \frac{2}{3}N$  modes,

left with

$$\frac{1}{2}\frac{d}{dt}\int (u_m - u)^2 dx = \frac{d}{dt}\int \left(\frac{|u_m|^2}{2} - \frac{|u|^2}{2} - u(u_m - u)\right) dx$$
$$= \frac{1}{2}\frac{d}{dt}\int |u_m|^2 dx - \int \partial_t (u(u_m - u)) dx =: \mathcal{I}_1 + \mathcal{I}_2.$$

Unlike the  $L^2$  conservation of the spectral solution  $u_N$ , consult (3.4), there is no  $L^2$ -energy conservation for the 2/3 smoothed solution  $u_m$  and we therefore leave  $\mathcal{I}_1$  is left as perfect time derivative. As for the second term

$$\mathcal{I}_2 = \int \partial_t \big( u(u_m - u) \big) dx \equiv \int \partial_t u(u_m - u) dx + \int u(\partial_t u_m - \partial_t u) dx,$$

we reproduce the same steps we had in the spectral case: using (3.1) and (3.8) to convert time derivatives to spatial ones, we find

$$\begin{aligned} \mathcal{I}_{2} &= -\int u u_{x}(u_{m}-u) dx - \int u \partial_{x} \left(\frac{u_{m}^{2}}{2} - \frac{u^{2}}{2}\right) dx + \int u \partial_{x}(I-\mathcal{S})[u_{m}^{2}] dx \\ &= -\int u u_{x}(u_{m}-u) dx + \int u_{x} \left(\frac{u_{m}^{2}}{2} - \frac{u^{2}}{2}\right) dx - \frac{1}{2} \int u_{x}(I-\mathcal{S})[u_{m}^{2}] dx \\ &= \int u_{x} \left(\frac{u_{m}^{2}}{2} - \frac{u^{2}}{2} - u(u_{m}-u)\right) dx - \frac{1}{2} \int u_{x}(I-\mathcal{S})[u_{m}^{2}] dx. \end{aligned}$$

Eventually, we end up with

$$\frac{1}{2}\frac{d}{dt}\int |u_m(x,t)-u(x,t)|^2 dx \leq \frac{|u_x|_{\infty}}{2}\int |u_m(x,t)-u(x,t)|^2 dx + \frac{1}{2}e_N(t) + \frac{1}{2}\frac{d}{dt}\int |u_m(x,t)|^2 dx,$$

where the error term,  $e_N$  is given by  $e_N(t) := -\int u_m^2 (I - S)[u_x] dx$ . Integrating in time we find

(3.10) 
$$\int_{x} |u_{m}(x,t) - u(x,t)|^{2} dx - \int_{x} |u_{m}(x,0) - u(x,0)|^{2} dx$$
$$\leq |u_{x}|_{\infty} \int_{\tau=0}^{t} \int |u_{m}(x,t) - u(x,t)|^{2} dx d\tau + \int_{0}^{t} e_{N}(\tau) d\tau + f_{N}(t),$$

with the additional error term,  $f_N(t)$ , given by

$$f_N(t) := \int |u_m(x,t)|^2 dx - \int |u_m(x,0)|^2 dx.$$

The error term  $e_N(t)$  can be estimated as before: observe that under the hypothesis  $u_x \in L^{\infty}_t C^{0,\alpha}_x$ , one has

(3.11) 
$$|e_N(t)| \lesssim \max_x |(I - \mathcal{S})[u_x(x, t)]| \cdot ||u_m||_{L^2}^2 \lesssim \frac{\ln N}{N^{\alpha}} ||u_N(\cdot, 0)||_{L^2}^2 \to 0.$$

To address the new error term,  $f_N(t)$ , we observe by the  $L^2_{\mathcal{S}}$ -energy conservation (3.9),

$$\int |u_m(x,t)|^2 dx = \sum_{|k| \le \frac{2}{3}N} \sigma_k^2 |\widehat{u}_k(t)|^2 \le \sum_{|k| \le \frac{2}{3}N} \sigma_k |\widehat{u}_k(t)|^2 = \sum_{|k| \le \frac{2}{3}N} \sigma_k |\widehat{u}_k(0)|^2$$
$$= \sum_{|k| \le \frac{2}{3}N} \sigma_k^2 |\widehat{u}_k(0)|^2 + (\sigma_k - \sigma_k^2) |\widehat{u}_k(0)|^2$$

C. BARDOS AND E. TADMOR

$$= \int |u_m(x,0)|^2 + \sum_{|k| \le \frac{2}{3}N} \left(\sigma_k - \sigma_k^2\right) |\widehat{u}_k(0)|^2.$$

Since  $\sigma_k \equiv 1$  for |k| < N/3, consult (3.6c), we conclude

$$(3.12) \quad f_N(t) := \int |u_m(x,t)|^2 - \int |u_m(x,0)|^2 \\ \leq \sum_{\frac{1}{3}N \le |k| \le \frac{2}{3}N} \left(\sigma_k - \sigma_k^2\right) |\widehat{u}_k(0)|^2 \le \left\| \left( P_{\frac{2}{3}N} - P_{\frac{1}{3}N} \right) u(\cdot,0) \right\|_{L^2}^2 \to 0$$

With (3.10), (3.12) and (3.11) in place, one obtains an estimate on the error integrated in space-time

$$\frac{1}{2}\frac{d}{dt}E_m(t) \le \frac{|u_x|_{\infty}}{2}E_m(t) + \frac{1}{2}\int_0^t e_N(\tau)d\tau + \frac{1}{2}f_N(t), \quad E_m(t) := \int_0^t \int |u_m(x,\tau) - u(x,\tau)|^2 dxd\tau$$

Convergence follows by Gronwall's inequality,

$$\int |u_m(x,t) - u(x,t)|^2 dx \lesssim e^{\int_0^t |u_x(\cdot,\tau)|_{\infty} d\tau} (\|u_m(\cdot,0) - u(\cdot,0)\|_{L^2}^2 + \max_{x,\tau \le t} |(I-\mathcal{S})u_x(x,\tau)| + \|(P_{\frac{2}{3}N} - P_{\frac{1}{3}N})u(\cdot,0)\|_{L^2}^2).$$

Moreover, with  $u_N(\cdot, 0) = P_N u(\cdot, 0)$  we end up with spectral convergence rate estimate

$$\int |u_m(x,t) - u(x,t)|^2 dx$$
  
$$\lesssim e^{\int_0^t |u_x(\cdot,\tau)|_\infty d\tau} \left( N^{-2s} ||u(\cdot,0)||^2_{H^s} + N^{\frac{3}{2}-s} \max_{\tau \le t} ||u(\cdot,\tau)||_{H^s} \right), \quad s > \frac{3}{2}.$$

## 4. Fourier method for Burgers equation: instability for weak solutions

In this section we discuss the spectral and the 2/3 de-aliased pseudo-spectral Fourier approximations of Burgers' equation, (3.1), after the formation of shock discontinuities. We show that both methods are unstable after the critical time,  $t > T_c$ . Recall that the spectral method is a special case of the 2/3 de-aliased method when we set the smoothing factors  $\sigma_k \equiv 1$ , see corollary 3.2. It will therefore suffice to consider the 2/3 de-aliasing pseudo-spectral Fourier method (3.8). We begin with its  $L_S^2$ -conservation (3.9), which we express as

(4.1) 
$$\|\mathcal{S}^{1/2}u_N(\cdot,t)\|_{L^2} = \|\mathcal{S}^{1/2}u_N(\cdot,0)\|_{L^2}, \qquad \mathcal{S}^{1/2}u_N := \sum_{|k| \le m} \sqrt{\sigma_k} \,\widehat{u}_k(t).$$

Since the quadratic energy associated with  $S^{1/2}u_N$  is bounded, it follows that, after extracting a subsequence if necessary<sup>4</sup> that  $S^{1/2}u_N(\cdot, t)$  and hence  $u_m = Su_N$  has a  $L^2$ -weak

<sup>&</sup>lt;sup>4</sup>Here and below we continue to label such subsequences as  $u_N$ .

limit,  $\overline{u}(x,t)$ . But  $\overline{u}$  cannot be the physically relevant entropy solution of (2.1). Our next result quantifies what can go wrong.

**Theorem 4.1 (The** 2/3 **method must admit spurious oscillations).** Let  $T_c$  be the critical time of shock formation in Burgers' equation (3.1). Let  $u_m = Su_N$  denote the smoothed 2/3 de-aliasing Fourier method, (3.6). Assume the  $L^6$ -bound,  $||u_m(\cdot,t)||_{L^6} \leq Const$  holds. Then, for  $t > T_c$ , there exists a constant  $c_0 > 0$  (independent of N) such that<sup>5</sup>

(4.2) 
$$\max_{x} |u_m(x,t)| \times ||u_m(\cdot,t)||_{TV}^2 \ge c_0 \sqrt{m}.$$

Theorem 4.1 implies that either the solution of the 2/3 de-aliasing Fourier method,  $u_m = Su_N$ , grows unboundedly,

$$\lim_{N \to \infty} \|u_m(\cdot, t)\|_{L^{\infty}} \longrightarrow \infty,$$

or it has an unbounded total variation of order  $\geq \mathcal{O}(\sqrt[4]{N})$ . Each one of these scenarios implies that  $u_m$  contains spurious oscillations which are noticeable *throughout* the computational domain, in agreement with the numerical evidence observed in [Tad89]. We note that this type of nonlinear instability applies to both, the 2/3 method and in particular, the spectral Fourier method and we refer in this context to the recent detailed study in [RFNM11, PNFS13] and the references therein.

*Proof.* We begin with (3.8)

(4.3) 
$$\frac{\partial}{\partial t}u_m(x,t) + \frac{1}{2}\frac{\partial}{\partial x}u_m^2(x,t) = \frac{1}{2}\frac{\partial}{\partial x}(I-\mathcal{S})[u_m^2](x,t).$$

Observe that the residual on the right tends to zero in  $H^{-1}$ ,

$$\begin{split} \left| \int \frac{\partial}{\partial x} \varphi(x) (I - \mathcal{S})[u_m^2](x, t) dx \right| &= \left| \int \left( (I - \mathcal{S}) \frac{\partial}{\partial x} \varphi(x) \right) u_m^2(x, t) dx \right| \\ &\leq \|u_m(\cdot, t)\|_{L^4}^2 \times \|(I - \mathcal{S}) \varphi_x(\cdot)\|_{L^2} \to 0, \qquad \forall \varphi \in H^1. \end{split}$$

Next, we consider the  $L^2$ -energy balance associated with (4.3). Multiplication by  $u_m$  yields

(4.4) 
$$\frac{1}{2}\frac{\partial}{\partial t}u_m^2(x,t) + \frac{1}{3}\frac{\partial}{\partial x}u_m^3(x,t) = \frac{1}{2}u_m(x,t)\frac{\partial}{\partial x}(I-\mathcal{S})[u_m^2](x,t).$$

We continue our argument by claiming that if (4.2) fails, then the energy production on the right of (4.4) also tends weakly to zero in  $H^{-1}$ . To this end, we examine the weak form of the expression on the right which we rewrite as

$$\int \varphi(x)u_m(x,t)\frac{\partial}{\partial x}(P_{2m}-\mathcal{S})[u_m^2](x,t)dx = \int (P_{2m}-\mathcal{S})\Big(\varphi(x)u_m(x,t)\Big)\frac{\partial}{\partial x}u_m^2(x,t)dx$$

It does not exceed

(4.5a)  

$$\begin{aligned} \left| \int \varphi(x) u_m \frac{\partial}{\partial x} (I - \mathcal{S}) [u_m^2](x, t) dx \right| \\ &= \left| \int (P_{2m} - \mathcal{S}) \Big( \varphi(x) u_m(x, t) \Big) u_m(x, t) \frac{\partial}{\partial x} u_m(x, t) dx \right| \\ &\leq \left\| (P_{2m} - \mathcal{S}) \big( \varphi(x) u_m(x, t) \big) \right\|_{L^{\infty}} \times \|u_m(\cdot, t)\|_{TV} \times \|u_m(\cdot, t)\|_{L^{\infty}} \end{aligned}$$

 $<sup>{}^{5}||</sup>u_{m}||_{TV}$  denotes the total variation of  $u_{m}$ .

To upper bound the first term we use standard decay estimate,  $|\sigma_j \hat{u}_N(j)(t)| \leq ||u_m(\cdot, t)||_{TV}/(1+|j|)$ . Noting that  $P_{2m} - S$  annihilates the first m/2 modes, namely, the multipliers  $\widehat{P_{2m}} - S(k) = 0$ ,  $|k| \leq m/2 = N/3$ , we find

$$\left| (P_{2m} - \mathcal{S}) \left( \varphi(x) u_m(x, t) \right) \right| _{L^{\infty}} \leq \sum_{\frac{m}{2} \le |k| \le 2m} (1 - \sigma_k) \left| \sum_{|j| \le m} \widehat{\varphi}(k - j) \sigma_j \widehat{u}_N(j, t) \right|$$

(4.5b)

$$\lesssim \sum_{\frac{m}{2} \le |k| \le 2m} \sqrt{\sum_{|j| \le m} (1 + |k - j|^2) |\widehat{\varphi}(k - j)|^2} \cdot \sqrt{\sum_{|j| \le m} \frac{1}{(1 + |k - j|^2)(1 + |j|^2)}} \times \|u_m(\cdot, t)\|_{TV}$$
  
 
$$\lesssim \|\varphi\|_{H^1} \|u_m(\cdot, t)\|_{TV} \times \frac{1}{\sqrt{m}}.$$

The last two inequalities (4.5) give us,

$$\left|\int \varphi(x)u_m(x,t)\frac{\partial}{\partial x}(I-\mathcal{S})[u_m^2](x,t)dx\right| \lesssim \frac{1}{\sqrt{m}} \|u_m(\cdot,t)\|_{TV}^2 \times \|u_m(\cdot,t)\|_{L^{\infty}} \times \|\varphi\|_{H^{1,2}}$$

We claim that (4.2) holds by contradiction. If it fails, then we can choose a subsequence,  $u_{m_k}$ , such that

$$\frac{1}{m_k} \|u_{m_k}(\cdot, t)\|_{TV}^2 \times \|u_{m_k}(\cdot, t)\|_{L^{\infty}} \le c_k, \qquad c_k \downarrow 0,$$

and the energy production on the right of (4.4) vanishes in  $H^{-1}$ . By assumption  $u_m^r(\cdot, t) \in L^2$  for r = 1, 2, 3 and the div-curl lemma, [Mu78, Tar79, Tar87] applies: it follows that  $\overline{u}$  is in fact a *strong*  $L^2$ -limit,  $u_{m_k} \to \overline{u}$ . Passing to the weak limit in (4.3)<sub>m\_k</sub> we have that  $\overline{u}$  is weak solution of Burgers' equation (3.1),

$$\frac{\partial}{\partial t}\overline{u}(x,t) + \frac{\partial}{\partial x}\left(\frac{\overline{u}^2(x,t)}{2}\right) = 0.$$

Moreover, passing to the weak limit in the energy balance  $(4.4)_{m_k}$ , we conclude that  $\overline{u}$  satisfies the quadratic entropy *equality* 

$$\frac{\partial}{\partial t} \left( \frac{\overline{u}^2(x,t)}{2} \right) + \frac{\partial}{\partial x} \left( \frac{\overline{u}^3(x,t)}{3} \right) = 0.$$

But, due to the uniqueness enforced with by the single entropy – in this case, the  $L^2$  energy, [Pan94], there exists no energy *conservative* weak solution of Burgers equation (3.1) after the critical time of shock formation.

**Remark 4.2.** The same result of instability holds if we employ the pseudo-spectral Fourier method with a general smoothing operator beyond just the 2/3 smoothing, namely  $Su_N = \sum_{|k| \leq N} \sigma_k \hat{u}_k e^{ikx}$  and smoothing factors  $\sigma_k$  decay too fast as  $|k| \uparrow N$ .

## 5. Fourier method for Euler equations: convergence for smooth solutions

Convergence of the spectral and pseudo-spectral approximation for the Burgers equation made use of its quadratic flux,  $u^2/2$ . The same approach can be pursued for the Euler equations,

(5.1) 
$$\frac{\partial}{\partial t}\mathbf{u} + \mathbb{P}\nabla_{\mathbf{x}}(\mathbf{u}\otimes\mathbf{u}) = 0, \qquad \mathbf{x}\in\mathbb{T}^d,$$

where  $\mathbb{P} := Id - \nabla_{\mathbf{x}} \Delta^{-1} \operatorname{div}_{\mathbf{x}}$  is the Leray projection into divergence free vector fields.

5.1. Convergence of spectral Fourier approximation for Euler equations. The spectral method for the Euler equations reads

(5.2) 
$$\frac{\partial}{\partial t}\mathbf{u}_N + \mathbb{P}\nabla_{\mathbf{x}}P_N(\mathbf{u}_N \otimes \mathbf{u}_N) = 0$$

Convergence for smooth solutions in this case, is in fact even simpler than in Burgers' equation. Observe that for any divergence free vectors fields,  $\mathbf{v}$  and  $\mathbf{u}$ , the following identity holds

$$\int \left\langle \left( \mathbf{v} \nabla_{\mathbf{x}} (\mathbf{v} \otimes \mathbf{v}) - \mathbf{v} \nabla_{\mathbf{x}} (\mathbf{u} \otimes \mathbf{u}) \right), \mathbf{v} - \mathbf{u} \right\rangle d\mathbf{x} \equiv \int \left\langle (\mathbf{v} - \mathbf{u}), \mathrm{S}[\mathbf{u}] (\mathbf{v} - \mathbf{u}) \right\rangle d\mathbf{x},$$

where  $S[\mathbf{u}]$  is the symmetric part of the stress tensor  $S[\mathbf{u}] := \frac{1}{2}(\nabla_{\mathbf{x}}\mathbf{u} + \nabla_{\mathbf{x}}\mathbf{u}^{\top})$ . We therefore have,

$$\left|\int \left\langle \mathbb{P}\nabla_{\mathbf{x}}(\mathbf{u}_N \otimes \mathbf{u}_N) - \mathbb{P}\nabla_{\mathbf{x}}(\mathbf{u} \otimes \mathbf{u}), (\mathbf{u}_N - \mathbf{u})\right\rangle d\mathbf{x} \right| \leq \|\nabla_{\mathbf{x}}\mathbf{u}\|_{L^{\infty}} \|\mathbf{u}_N - \mathbf{u}\|_{L^2}^2,$$

The error equation

$$\frac{\partial}{\partial t}(\mathbf{u}_N - \mathbf{u}) + \mathbb{P}\nabla_{\mathbf{x}}(\mathbf{u}_N \otimes \mathbf{u}_N) - \mathbb{P}\nabla_{\mathbf{x}}(\mathbf{u} \otimes \mathbf{u}) = (I - P_N)\mathbb{P}\nabla_{\mathbf{x}}(\mathbf{u}_N \otimes \mathbf{u}_N),$$

implies

(5.3) 
$$\frac{1}{2} \frac{d}{dt} \|\mathbf{u}_N - \mathbf{u}\|_{L^2}^2$$
$$\leq \|\nabla_{\mathbf{x}} \mathbf{u}\|_{L^{\infty}} \|\mathbf{u}_N - \mathbf{u}\|_{L^2}^2 + \left| \int \langle (I - P_N) [\mathbb{P}\nabla_{\mathbf{x}} (\mathbf{u}_N \otimes \mathbf{u}_N)], \mathbf{u}_N - \mathbf{u} \rangle \, d\mathbf{x} \right|$$
$$\leq \|\nabla_{\mathbf{x}} \mathbf{u}\|_{L^{\infty}} \|\mathbf{u}_N - \mathbf{u}\|_{L^2}^2 + \left| \int \langle ((I - P_N)\nabla \mathbf{u}) \, \mathbf{u}_N, \mathbf{u}_N \rangle \, d\mathbf{x} \right|.$$

Arguing along the lines of our convergence statement for Burgers equations we conclude that the following result holds.

Theorem 5.1 (Spectral convergence for smooth solutions of Euler equations). Assume that for  $0 < t < T_c$ , the solution of the Euler equations (5.1) is smooth,  $\mathbf{u}(\cdot, t) \in L^{\infty}([0, T_c), C^{1+\alpha}(0, 2\pi])$ . Then its spectral Fourier approximation (5.2) converges in  $L^{\infty}([0, T_c], L^2(\mathbb{T}^d))$ ,

$$\|\mathbf{u}_N(\cdot, t) - \mathbf{u}(\cdot, t)\|_{L^2} \to 0, \qquad 0 \le t < T_c,$$

and the following spectral convergence rate estimate holds

$$\begin{aligned} \|\mathbf{u}_{N}(\cdot,t) - \mathbf{u}(\cdot,t)\|^{2} \\ &\lesssim e^{2\int_{0}^{t} |\nabla \mathbf{u}(\cdot,\tau)|_{\infty} d\tau} \left( N^{-2s} \|\mathbf{u}(\cdot,0)\|_{H^{s}}^{2} + N^{\frac{d}{2}+1-s} \max_{\tau \leq t} \|\mathbf{u}(\cdot,\tau)\|_{H^{s}} \right), \quad s > \frac{d}{2} + 1. \end{aligned}$$

*Proof.* Integrating (5.2) against  $\mathbf{u}_N$  we find the usual statement of  $L^2$  energy conservation,

$$\|\mathbf{u}_N(\cdot,t)\|_{L^2}^2 = \|\mathbf{u}_N(\cdot,0)\|_{L^2}^2.$$

Using (5.3), we conclude

 $\|\mathbf{u}_N(\cdot,t) - \mathbf{u}(\cdot,t)\|_{L^2}^2 \lesssim e^{2U'_{\infty}(t;0)} \|(I - P_N)\mathbf{u}(\cdot,0)\|_{L^2}^2$ 

$$+\|\mathbf{u}_{N}(\cdot,0)\|_{L^{2}}^{2}\int_{0}^{t}e^{2U_{\infty}'(t;\tau)}\|(I-P_{N})\nabla\mathbf{u}(\cdot,\tau)\|_{L^{\infty}}d\tau, \quad U_{\infty}'(t;\tau):=\int_{s=\tau}^{t}\|\nabla_{\mathbf{x}}\mathbf{u}(\cdot,s)\|_{L^{\infty}}ds,$$

which yields the spectral convergence rate estimate

(5.4) 
$$\|\mathbf{u}_N(\cdot,t) - \mathbf{u}(\cdot,t)\|_{L^2}^2$$
  
 $\lesssim e^{2U'_{\infty}(t;0)} \left( N^{-2s} \|\mathbf{u}(\cdot,0)\|_{H^s}^2 + N^{-s+\frac{d}{2}+1} \max_{\tau \le t} \|\mathbf{u}(\cdot,\tau)\|_{H^s} \right), \qquad s > \frac{d}{2} + 1.$ 

Observe that the error estimate in the case of Euler equation depends on the truncation error of  $\nabla_{\mathbf{x}} \mathbf{u}$ , corresponding to the dependence on the truncation error of  $u_x$  in Burgers equation. The additional loss factor of d/2 is due to the  $L^{\infty}(\mathbb{T}^d)$ -bound,  $\max_{\mathbf{x}} |(I-P_N)\mathbf{w}(\mathbf{x})| \leq ||\mathbf{w}||_{H^s}$  for s > d/2, consult (1.2).

5.2. The 2/3 pseudo-spectral approximation of Euler equations. The pseudo-spectral Fourier method for the Euler equations reads

$$\frac{\partial}{\partial t}\mathbf{u}_N + \mathbb{P}\nabla_{\mathbf{x}}\psi_N(\mathbf{u}_N\otimes\mathbf{u}_N) = 0,$$

Observe that since  $\psi_N$  does not commute with  $\mathbb{P}\nabla_{\mathbf{x}}$ , there is no  $L^2$ -energy conservation. We introduce the smoothing operator  $\mathcal{S}\mathbf{u}_N := \sum_{|k| \leq m} \sigma_k \widehat{\mathbf{u}}_k(t)$  which acts on wavenumbers  $|k| \leq m = \frac{2}{3}N$ , while leaving the first 1/3 portion of the spectrum unchanged:  $\sigma_k = \sigma(|k|/N)$ , where  $\sigma(1-\sigma)$  is supported in  $(\frac{1}{3}, \frac{2}{3})$ . The resulting 2/3 de-aliasing pseudo-spectral method reads

(5.5) 
$$\frac{\partial}{\partial t}\mathbf{u}_N + \mathbb{P}\nabla_{\mathbf{x}}\psi_N(\mathcal{S}\mathbf{u}_N \otimes \mathcal{S}\mathbf{u}_N) = 0.$$

It is the 2/3 Fourier method which is being used in actual computations, e.g., [OHFS10, KH89, Kerr93, Kerr05] and the references therein. Next, we act with the smoothing S: arguing along the lines of the 2/3 method for the Burgers' equation in corollary 3.2, we find that the  $\mathbf{u}_m := S \mathbf{u}_N$  satisfies the aliasing-free equation

(5.6) 
$$\frac{\partial}{\partial t}\mathbf{u}_m + \mathcal{SP}\nabla_{\mathbf{x}}(\mathbf{u}_m \otimes \mathbf{u}_m) = 0.$$

Observe that since  $\mathcal{S}$  commutes with differentiation,  $\mathbf{u}_m$  retains incompressibility,

$$\frac{\partial}{\partial t}\mathbf{u}_m + \mathbb{P}\nabla_{\mathbf{x}}\mathcal{S}(\mathbf{u}_m \otimes \mathbf{u}_m) = 0.$$

As before, we can integrate against  $\mathbf{u}_N$  to find by incompressibility of  $\mathbf{u}_m$ ,

$$\frac{1}{2}\frac{d}{dt}\int \langle \mathbf{u}_N(\mathbf{x},t), \mathbf{u}_m(\mathbf{x},t) \rangle d\mathbf{x} = -\int \langle \mathcal{S}\mathbf{u}_N, \mathbb{P}\nabla_{\mathbf{x}}(\mathbf{u}_m \otimes \mathbf{u}_m) \rangle d\mathbf{x} = 0,$$

which implies the weighted  $L_{S}^{2}$ -energy conservation,

(5.7) 
$$\|\mathbf{u}_N(\cdot,t)\|_{L^2_{\mathcal{S}}}^2 = \|\mathbf{u}_N(\cdot,0)\|_{L^2_{\mathcal{S}}}^2, \qquad \|\mathbf{u}_N(\cdot,t)\|_{L^2_{\mathcal{S}}}^2 := (2\pi)^d \sum \sigma_{\mathbf{k}} |\widehat{\mathbf{u}}_{\mathbf{k}}(t)|^2.$$

**Theorem 5.2** (Spectral convergence of 2/3 method for smooth Euler solutions). Assume that for  $0 < t < T_c$ , the solution of the Euler equations (5.1) is smooth,  $\mathbf{u}(\cdot, t) \in L^{\infty}([0, T_c), C^{1+\alpha}(0, 2\pi])$ . Then, the smoothed solution  $\mathbf{u}_m = S\mathbf{u}_N$  of its 2/3 de-aliasing pseudo-spectral Fourier approximation (5.5) converges in  $L^{\infty}([0, T_c], L^2(\mathbb{T}^d))$ ,

$$\|\mathbf{u}_m(\cdot, t) - \mathbf{u}(\cdot, t)\|_{L^2} \to 0, \qquad 0 \le t < T_c,$$

20

and the following spectral convergence rate estimate holds

$$\begin{split} \|\mathbf{u}_{m}(\cdot,t) - \mathbf{u}(\cdot,t)\|_{L^{2}}^{2} \\ \lesssim e^{2 \int_{0}^{t} |\nabla \mathbf{u}(\cdot,\tau)|_{\infty} d\tau} \left( N^{-2s} \|\mathbf{u}(\cdot,0)\|_{H^{s}}^{2} + N^{\frac{d}{2}+1-s} \max_{\tau \leq t} \|\mathbf{u}(\cdot,\tau)\|_{H^{s}} \right), \quad s > \frac{d}{2} + 1. \end{split}$$

*Proof.* We rewrite (5.6) in the form

$$\frac{\partial}{\partial t}\mathbf{u}_m + \mathbb{P}\nabla_{\mathbf{x}}(\mathbf{u}_m \otimes \mathbf{u}_m) = (I - \mathcal{S})\left(\mathbb{P}\nabla_{\mathbf{x}}(\mathbf{u}_m \otimes \mathbf{u}_m)\right).$$

Subtract the exact equation (5.1): using the identity (5.3) we find, as before

(5.8) 
$$\frac{1}{2} \frac{d}{dt} \|\mathbf{u}_{m} - \mathbf{u}\|_{L^{2}}^{2}$$
$$\leq \|\nabla_{\mathbf{x}}\mathbf{u}\|_{L^{\infty}} \|\mathbf{u}_{m} - \mathbf{u}\|_{L^{2}}^{2} + \left|\int \langle (I - \mathcal{S})[\mathbb{P}\nabla_{\mathbf{x}}(\mathbf{u}_{m} \otimes \mathbf{u}_{m})], \mathbf{u}_{m} - \mathbf{u} \rangle d\mathbf{x}\right|$$
$$\leq \|\nabla_{\mathbf{x}}\mathbf{u}\|_{L^{\infty}} \|\mathbf{u}_{m} - \mathbf{u}\|_{L^{2}}^{2} + \left|\int \langle ((I - \mathcal{S})\nabla_{\mathbf{x}}\mathbf{u}) \mathbf{u}_{m}, \mathbf{u}_{m} \rangle d\mathbf{x}\right|$$
$$+ \left|\int \langle ((I - \mathcal{S})\nabla_{\mathbf{x}}\mathbf{u}_{m}) \mathbf{u}_{m}, \mathbf{u}_{m} \rangle d\mathbf{x}\right|.$$

The last term on the right is due to the fact that (I - S) need not annihilate  $\nabla_{\mathbf{x}} \mathbf{u}_m$ . However, since  $\mathbf{u}_m$  is incompressible, we find

$$\int \langle ((I - S) \nabla_{\mathbf{x}} \mathbf{u}_m) \mathbf{u}_m, \mathbf{u}_m \rangle d\mathbf{x} = \sum_{\alpha, \beta} \int \mathbf{u}_{m\alpha} \partial_\alpha \mathbf{u}_{m\beta} (I - S) \mathbf{u}_{m\beta} d\mathbf{x}$$
$$= \sum_{\alpha, \beta} \int \mathbf{u}_{m\alpha} \frac{1}{2} \partial_\alpha \left( \mathbf{u}_{m\beta} (I - S) \mathbf{u}_{m\beta} \right) d\mathbf{x}$$
$$= -\frac{1}{2} \int \sum_{\alpha} \partial_\alpha \mathbf{u}_{m\alpha} \sum_{\beta} \left( \mathbf{u}_{m\beta} (I - S) \mathbf{u}_{m\beta} \right) d\mathbf{x} = 0$$

We end up with the error bound

$$\begin{aligned} \|\mathbf{u}_{m}(\cdot,t) - \mathbf{u}(\cdot,t)\|_{L^{2}}^{2} &\lesssim e^{2U_{\infty}'(t;0)} \|(I-\mathcal{S})\mathbf{u}(\cdot,0)\|_{L^{2}}^{2} \\ &+ \|\mathbf{u}_{m}(\cdot,0)\|_{L^{2}_{\mathcal{S}}}^{2} \int_{0}^{t} e^{2U_{\infty}'(t;\tau)} \|(I-\mathcal{S})\nabla_{\mathbf{x}}\mathbf{u}(\cdot,\tau)\|_{L^{\infty}} d\tau, \quad U_{\infty}'(t;\tau) := \int_{s=\tau}^{t} \|\nabla_{\mathbf{x}}\mathbf{u}(\cdot,s)\|_{L^{\infty}} ds, \end{aligned}$$
and spectral convergence rate follows.

and spectral convergence rate follows.

# 6. Fourier method for Euler equations: failure of convergence for weak SOLUTIONS?

We now consider the convergence of the 2/3 method (5.5) for weak solutions of Euler equations. Its *m*-mode de-aliased solution is governed by (5.6)

(6.1) 
$$\frac{\partial}{\partial t}\mathbf{u}_m + \mathcal{SP}\nabla_{\mathbf{x}}\left(\mathbf{u}_m \otimes \mathbf{u}_m\right) = 0.$$

The method is energy preserving in the sense that  $S^{1/2}\mathbf{u}_N$  is  $L^2$ -conservative, (5.7), and hence  $\mathbf{u}_m = S\mathbf{u}_N$  has s a weak limit,  $\overline{\mathbf{u}}$ . The question is to characterize whether  $\overline{\mathbf{u}}(x,t)$  is an energy conserving weak solution of Euler equations (5.1),

(6.2) 
$$\frac{\partial}{\partial t}\overline{\mathbf{u}} + \mathbb{P}\nabla_{\mathbf{x}}(\overline{\mathbf{u}}\otimes\overline{\mathbf{u}}) = 0$$

To this end we compare (5.5) and (6.2): since  $\mathbf{u}_m$  tends weakly to  $\overline{\mathbf{u}}$  and  $\partial_t \mathbf{u}_m \to \partial_t \overline{\mathbf{u}}$ , then comparing the remaining spatial parts of (5.5) and (6.2), yields that  $S\mathbb{P}[\mathbf{u}_m \otimes \mathbf{u}_m](x,t)$  and hence  $\mathbb{P}[\mathbf{u}_m \otimes \mathbf{u}_m](x,t)$  tends weakly to  $\mathbb{P}[\overline{\mathbf{u}} \otimes \overline{\mathbf{u}}](x,t)$ . This, however, is not enough to imply the strong convergence of  $u_N$ , as shown by a simple counterexample of a 2D potential flow,  $\mathbf{u}_n = \nabla_{\mathbf{x}}^{\perp} \Phi_n$  where

$$\Phi_n(x_1, x_2) = \frac{1}{n} \Xi(x_1, x_2)(\sin nx_1 + \sin nx_2)$$

with  $\Xi(x_1, x_2) \in \mathcal{D}(\mathbb{R}^2)$  localized near any point (say (0, 0)) with weak limit  $\overline{\mathbf{u}} \equiv 0$ . In this case  $w - \lim_{N \to \infty} \nabla \mathbb{P}(u_N \otimes u_N) = \nabla \mathbb{P}(\overline{\mathbf{u}} \otimes \overline{\mathbf{u}}) = 0$ , yet

$$w \lim_{N \to \infty} (u_{N1})^2 = w \lim_{N \to \infty} (u_{N2})^2 = \frac{\Xi(x_1, x_2)^2}{2} \neq 0.$$

Although  $\overline{\mathbf{u}}$  need not be a weak solution of Euler equations, it satisfies a weaker notion of a *dissipative solution* in the sense of DiPerna-Lions [Lions96] To this end, let  $\mathbf{w}$  a divergence-free smooth solution of

(6.3) 
$$\partial_t \mathbf{w} + \mathbb{P}(\nabla \mathbf{w} \otimes \mathbf{w}) = E(\mathbf{w}), \qquad \mathbb{P}E(\mathbf{w}) = 0.$$

Now, compare it with the 2/3 solution (6.1): the same computation with Gronwall lemma leads to,

$$\begin{aligned} \|(\mathbf{u}_{N} - \mathbf{w})(\cdot, t)\|_{L^{2}(\Omega)}^{2} &\leq e^{2W_{\infty}'(t;0)} \|(\mathbf{u}_{N} - \mathbf{w})(\cdot, 0)\|_{L^{2}(\Omega)}^{2} \\ &+ 2\|\mathbf{u}_{N}(\cdot, 0)\|_{L^{2}(\Omega)}^{2} \int_{0}^{t} \|(P_{N}\mathbf{w} - \mathbf{w}))(\cdot, s)\|_{W^{1,\infty}(\Omega)} \\ &+ 2\int_{0}^{t} e^{2W_{\infty}'(t;\tau)} \|(E(\mathbf{w}(\tau)), \mathbf{u}_{N}(\tau) - \mathbf{w}(\tau))\| d\tau, \quad W_{\infty}'(t;\tau) := \int_{s=\tau}^{t} \|\nabla_{\mathbf{x}}\mathbf{w}(\cdot, s)\|_{L^{\infty}(\Omega)} ds. \end{aligned}$$

Passing to the weak limit it follows that  $\overline{\mathbf{u}}$  is a dissipative solution, satisfying for all divergence-free smooth solution of (6.3), the stability estimate

$$\begin{split} \|(\overline{\mathbf{u}} - \mathbf{w})(\cdot, t)\|_{L^2(\Omega)}^2 &\leq e^{2W'_{\infty}(t;0)} \|(\overline{\mathbf{u}} - \mathbf{w})(\cdot, 0)\|_{L^2(\Omega)}^2 \\ &+ 2\int_0^t e^{2W'_{\infty}(t;\tau)} |(E(\mathbf{w}(\tau)), \overline{\mathbf{u}}(\tau) - \mathbf{w}(\tau)))| d\tau. \end{split}$$

The notion of dissipative solution can be instrumental in the context of stability near a smooth solution, **w**, or even in the context of uniqueness. However, the construction of [DLSz10] does not exclude the existence of rough initial data for which the Cauchy problem associated with Euler equations (5.1) have an infinite set of dissipative solutions. In fact, it is observed in [DLSz10] that any weak solution with a non-increasing energy,  $\|\mathbf{u}(\cdot, t)\|_{L^2} \leq \|\mathbf{u}(\cdot, 0)\|_{L^2}$ , is a dissipative solution. These, so-called admissible solutions, arise as solutions of the Cauchy problem for an infinite set of (rough) initial data, and can be obtained as strong limit in  $C(0, T; L^2_{weak}(\Omega))$  of solutions for the problem

$$\partial_t \mathbf{u}_N + \mathbb{P}(\nabla_{\mathbf{x}}(\mathbf{u}_N \otimes \mathbf{u}_N) = E_N)$$

with w-lim $E_N = 0$ , while  $\int \langle E_N, \mathbf{u}_N \rangle dx$  does not converge to 0.

We summarize the above observations, by stating that as long as the solution of the Euler equations remains sufficiently smooth, then its spectral and de-aliased pseudo-spectral approximations converge in  $L^2(\Omega)$ . Indeed, in theorems 5.1 and 5.2, we quantified the convergence rate for  $H^s$ -regular solutions **u**. If **u** has a minimal smoothness such that the vorticity  $\boldsymbol{\omega}_N := \nabla \times \mathbf{u}_N$  is compactly embedded in  $C([0,T], H^{-1}(\mathbb{R}^N))$ , then by the div-curl lemma,  $\mathbf{u}_N(\cdot, t)$  converges strongly in  $L^{\infty}([0,T], L^2_{loc}(\mathbb{R}^N))$  to an energy-preserving limit solution  $\overline{\mathbf{u}}$ , [LNT00].

The situation is different, however, when dealing with "rough" solutions of the underlying Euler equations. In the absence of any information re:the smoothness of the underlying Euler solutions (— as loss of smoothness for the 3D Euler equations is still a challenging open problem), energy-preserving numerical method need not shed light on the question of global regularity vs. finite-time blow-up. Recall that  $L^2$ -energy conservation was conjectured by Onsager [ON49] and verified in [Ey94, CET94, BT10] under the assumption of minimal smoothness of **u**, but otherwise is not supported by the energy decreasing solutions of Euler equation, [Co07, DeLS12, Buck13].

The similar scenario of quadratic entropy conservation in the context of Burgers' equations, is responsible for spurious oscillations, and its detailed analysis can be found in [La86] after [vN63]. Here, enforcing energy conservation at the "critical" time when Euler solutions seem to lose sufficient smoothness leads to nonlinear instability which manifests itself through oscillations noticeable throughout the computational domain, in agreement with the numerical evidence observed in [HL07], see Figure 6.4(a) below. The precise large-time behavior of the (pseudo-) spectral approximations is intimately related to a proper albeit yet unclear notion of propagating smoothness for solutions of Euler equations which, even if they do not explicitly blow up, may exhibit spurious oscillations due to the amplification factor in higher norms.

# 7. The spectral viscosity method: Nonlinear stability and spectral convergence

The nonlinear instability results in sections 4 and 6 emphasize the competition between spectral convergence for smooth solutions vs. nonlinear instabilities for problems which lack sufficient smoothness. One class of methods for nonlinear evolution equations which entertain both — spectral convergence and nonlinear stability, is the class *spectral viscosity* (SV) methods, introduced in [Tad89]. We demonstrate the SV method in the context of Burgers equation,

(7.1a) 
$$\frac{\partial}{\partial t}u_N(x,t) + \frac{1}{2}\frac{\partial}{\partial x}\left(\psi_N\left[u_N^2\right](x,t)\right) = SV[u_N](x,t), \qquad x \in \mathbb{T}([0,2\pi)).$$

On the right of (7.1a) we have added a judicious amount of spectral viscosity of order 2r:

(7.1b) 
$$SV[u_N](x,t) := -N \sum_{|k| \le N} \sigma\left(\frac{|k|}{N}\right) \widehat{u}_k(t) e^{ikx}, \quad \sigma(\xi) \lesssim \left(|\xi|^{2r} - \frac{1}{N}\right)_+, \qquad r \ge 1$$

Without it, the pseudo-spectral solution will develops spurious Gibbs oscillations after the formation of shocks. Observe that the spectral viscosity term in (7.1b) adds a *spectrally* 



FIGURE 6.4. A comparison of axial vorticity contours of 3D Euler computation [HL07] at t = 18 (top) and at t = 19 (bottom). On left(a): the solution obtained by the energy preserving  $\frac{2}{3}$  de-aliasing method; on right (b): the solution obtained by an energy decreasing smoothing of the Fourier method. The resolution is  $1024 \times 768 \times 2048$ .

small amount of numerical dissipation for high modes,  $k \gg 1$  (in contrast for "standard" finite-order amount of numerical dissipation in finite-difference methods),

$$\|SV[w]\|_{\dot{H}^{\alpha}} \lesssim N^{1-(\alpha-\beta)(1-\frac{1}{2r})} \|w\|_{\dot{H}^{\beta}}, \qquad \forall \beta \ll \alpha - 1 \in \mathbb{R}.$$

Indeed, the low-pass SV filter on the right of (7.1a) vanishes for modes  $|k| \leq N^{(2r-1)/2r}$ , which in turn leads to spectral convergence for smooth solutions. Arguing along the lines of theorem 3.3 we state the following.

Theorem 7.1 (Spectral convergence for smooth solutions of Burgers' equations). Consider the Burgers equation, (3.1), with a smooth solution  $u(\cdot, t) \in L^{\infty}([0, T_c], C^{1+\alpha}(0, 2\pi])$ . Then its spectral viscosity approximation (7.1),

$$\frac{d}{dt}u_N(x_\nu,t) + \frac{1}{2}\frac{\partial}{\partial x}\left(\psi_N\left[u_N^2\right](x,t)\right)\Big|_{x=x_\nu} = SV[u_N](x_\nu,t), \qquad \nu = 0, 1, \dots, 2N$$

converges,  $\|u_N(\cdot,t) - u(\cdot,t)\|_{L^2} \to 0$  for  $0 \le t \le T_c$  and the following spectral convergence rate estimate holds for all  $s > \frac{3}{2}$ ,

$$\begin{aligned} \|u_N(\cdot,t) - u(\cdot,t)\|^2 \\ \lesssim e^{\int_0^t |u_x(\cdot,\tau)|_\infty d\tau} \left( N^{-2s} \|u(\cdot,0)\|_{H^s}^2 + N^{\frac{2r-1}{2r}(\frac{3}{2}-s)} \max_{\tau \le t} \|u(\cdot,\tau)\|_{H^s} \right), \quad s > \frac{3}{2} \end{aligned}$$

At the same time, spectral viscosity is strong enough to enforce a sufficient amount of  $L^2$  energy dissipation, which in turn implies convergence after the formation of shock discontinuities. We quote below the convergence statement of the hyper-SV method. **Theorem 7.2 (Convergence of the hyper-SV method for Burgers equation** [Tad89, Tad93b, Tad04]). Let u be the unique entropy solution of the inviscid Burgers equation, (3.1), subject to uniformly bounded initial data  $u_0$ , and let  $u_N$  be the spectral viscosity approximation (7.1) subject to  $L^{\infty}$  data  $u_N(0) \approx u_0$ . Then, if  $u_N$  remains uniformly bounded<sup>6</sup> it converges to the unique entropy solution,  $||u_N(\cdot, t) - u(\cdot, t)||_{L^2} \to 0$ .

**Remark 7.3.** We note that unlike the 2/3 de-aliasing method, the SV method does *not* completely remove the high-frequencies but instead, it introduces "just the right amount" of smoothing for  $|k| \gg 1$  which enables to balance spectral accuracy with nonlinear stability. The SV method can be viewed as a proper smoothing which addresses the instability of general smoothing of the pseudo-spectral Fourier method sought in remark 4.2. Moreover, even after the formation of shock discontinuities, the SV solution still contains highly accurate information of the exact entropy solution which can be extracted by post-processing, [SW95].

Similar results of spectral convergence of SV methods hold in the context of incompressible Euler equations, [KK00, SS07, AX09],

(7.2) 
$$\frac{\partial}{\partial t} \mathbf{u}_N + \mathbb{P} \nabla_{\mathbf{x}} \psi_N \left( \mathcal{S} \mathbf{u}_N \otimes \mathcal{S} \mathbf{u}_N \right) = SV[\mathbf{u}_N],$$
$$SV[\mathbf{u}_N](\mathbf{x}, t) := -N \sum_{|\mathbf{k}| \le N} \sigma\left(\frac{|\mathbf{k}|}{N}\right) \widehat{\mathbf{u}}_{\mathbf{k}}(t) e^{i\mathbf{k} \cdot \mathbf{x}}.$$

In contrast to the spurious oscillations with the 2/3 methods shown in figure 6.4(a), the oscillations-free results in 6.4(b) correspond to the proper amount of smoothing employed in [HL07]. Thus, the issue of adding "just the right amount" of hyper-viscosity is particularly relevant in this context of Large Eddy Simulation (LES) for highly turbulent flows, when one needs to strike a balance between a sufficient amount of numerical dissipation for stability without giving up on high-order accuracy for physically relevant Euler (and Navier-Stokes solutions). The SV method in (7.2) adds this balanced amount of hyper-viscoisty, [KK00, GP03, SK04, SS07, PSSBS07].

#### 8. BEYOND QUADRATIC NONLINEARITIES: 1D ISENTROPIC EQUATIONS

We consider the one-dimensional isentropic equations in Lagrangian coordinate,

(8.1a) 
$$\frac{\partial}{\partial t}u + \frac{\partial}{\partial x}q(v) = 0, \quad q'(v) > 0$$

(8.1b) 
$$\frac{\partial}{\partial t}v + \frac{\partial}{\partial x}u = 0,$$

which is approximated by the spectral method

(8.2a) 
$$\frac{\partial}{\partial t}u_N + \frac{\partial}{\partial x}q(v_N) = (I - P_N)q(v_N),$$

(8.2b) 
$$\frac{\partial}{\partial t}v_N + \frac{\partial}{\partial x}u_N = 0.$$

Denote by U the vector of conservative variables,  $U := (u, v)^{\top}$ , by F(U) the corresponding flux,  $F(U) := (q(v), u)^{\top}$  and let  $\eta(U)$  be the entropy  $\eta(U) := \frac{1}{2}|u|^2 + Q(v), \ Q'(v) = q(v).$ 

<sup>&</sup>lt;sup>6</sup>The question of uniform boundedness of  $u_N$  was proved for the second order SV method, corresponding to r = 1, in [Tad93a], but it remains open for the hyper SV case with r > 1.

Multiplying the system by  $\nabla_U \eta(U)$  and integrating gives:

$$\frac{d}{dt}\int \left(\frac{|u_N|^2}{2} + u_N\partial_x q(v_N) + q(v_N)\partial_x u_N\right)dx = \int (I - P_N)q(v_N)u_Ndx = 0$$

and hence there the total entropy is conserved for both the exact an approximate solutions<sup>7</sup>

$$\partial_t \int \eta(U) dx = 0$$
 and  $\partial_t \int \eta(U_N) dx = 0.$ 

Continuing as in DiPerna-Chen [DiP83, Ch97, Daf79], we write

$$\partial_t \int \left( \eta(U_N) - \eta(U) - \left\langle \eta'(U), U_n - U \right\rangle \right) dx =$$

$$(8.3) \qquad \int \left\langle \eta''(U)U_t, (U_N - U) \right\rangle dx - \int \left\langle \eta'(U), (U_N)_t - U_t \right\rangle dx =$$

$$- \int \left\langle \eta''(U)F(U)_x, U_N - U \right\rangle dx - \int \left\langle \eta'(U), F(U_N) \right\rangle_x - F(U)_x \right\rangle dx + \text{error term}$$

$$=: \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3$$

The first two terms on the right amount to

$$\begin{aligned} |\mathcal{I}_{1}+\mathcal{I}_{2}| &= \left| \int \left\langle \eta''(U)F(U)_{x}, U_{N}-U \right\rangle dx + \int \left\langle \eta'(U), F(U_{N})_{x}-F(U)_{x} \right\rangle dx \right| \\ &= \left| \int \left\langle \eta''(U)F'(U)U_{x}, U_{N}-U \right\rangle dx - \left\langle \eta''(U)U_{x}, F(U_{N})-F(U) \right\rangle dx \right| \\ &= \left| \int \left\langle \eta''(U)F'(U)U_{x}, U_{N}-U \right\rangle dx - \left\langle \eta''(U)U_{x}, F'(U)U_{x}+\mathcal{O} \| U_{N}-U \|^{2} \right\rangle dx \right|. \end{aligned}$$

Since the entropy Hessian symmetrize the system, one has  $\eta''(U)F'(U) = F'(U)\eta''(U)$ , and we conclude that the last expression does not exceed

$$|\mathcal{I}_1 + \mathcal{I}_2| = \left| \int \left\langle \eta''(U) F'(U) U_x, U_N - U \right\rangle dx - \left\langle \eta''(U) U_x, F(U_N) - F(U) \right\rangle dx \right| \lesssim \|U\|_{C^1} \|U_N - U\|^2$$

On the other hand

$$\mathcal{I}_3 = \text{error term} = \int (I - P_N) q_x(v_N) (u - u_N) dx = \int \partial_x q(v_N) (I - P_N) u_x dx$$

which goes to zero for sufficiently smooth  $u \in C^{1+\alpha}$ . Inserting the last two bound into (8.3) we find that

$$\partial_t \int \left( \eta(U_N) - \eta(U) - \left\langle \eta'(U), U_n - U \right\rangle \right) dx \lesssim \|U\|_{C^1} \|U_N - U\|^2 + o(1).$$

By strict convexity, the integrand on the left is of order  $\sim ||U_N - U||^2$  and we conclude the following.

**Theorem 8.1.** Assume that for  $0 < t < T_c$ , the solution of the isentropic Euler equations (8.1) is smooth,  $U(\cdot, t) \in L^{\infty}([0, T_c), C^{1+\alpha}(0, 2\pi])$ . Then, its spectral approximation (8.2) converge in  $L_t^{\infty} L_x^2$ ,

$$||U_N(\cdot, t) - U(\cdot, t)||_{L^2} \to 0, \qquad 0 \le t < T_c.$$

<sup>&</sup>lt;sup>7</sup>This intriguing property seems specific to the isentropic equation in Lagrangian coordinate.

#### References

- [AGT86] S. Abarbanel, D. Gottlieb and E. Tadmor, Spectral methods for discontinuous problems "Numerical Methods for Fluid Dynamics II", Proceedings of the 1985 Conference on Numerical Methods for Fluid Dynamics (K. W. Morton and M. J. Baines, eds.), Clarendon Press, Oxford (1986), 129-153.
- [AX09] J. Avrin and C. Xiao, Convergence of Galerkin solutions and continuous dependence on data in spectrally hyperviscous models of 3D turbulent flow J. Differential Eqs 247 (2009) 2778-2798.
- [BT10] C. Bardos and E. Titi, Loss of smoothness and energy conserving rough weak solutions for the 3D Euler equations, DCDS Ser. 5, v 3(2) (2010), 185-197.
- [BT13] C. Bardos and E. Titi, Mathematics and Turbulence: where do we stand? Journal of Turbulence Volume 14, Issue 3, March 2013, pages 42-76.
- [Buck13] T. Buckmaster, Onsager's conjecture almost everywhere in time arXiv:1304.1049, 2013.
- [DLSz10] C. De Lellis and L. Székelyhidi, Jr., On admissibility criteria for weak solutions of the Euler equations Arch. Ration. Mech. Anal., 195 (1) (2010), 225–260.
- [Ch97] G.-Q Chen, Remarks on DiPerna's paper "Convergence of the viscosity method for isentropic gas dynamics", Proc. AMS 125(10), 1997, 2981-2986.
- [Co07] P. Constantin, On the Euler equations of incompressible fluids, Bull. AMS 44(4), (2007) 603-621.
- [CET94] P. Constantin, W. E and E. S. Titi, Onsager's conjecture on the Energy Conservation for Solutions of Euler's equation, Comm. Math. Phys., 165 (1994), 207-209.
- [Daf79] C. Dafermos, The second law of thermodynamics and stability, Arch. Rational Mech. Anal. 70 (199), 167-179.
- [DeLS12] C. De Lellis and L. Székelyhidi, The h-principle and the equations of fluid dynamics Bull. Amer. Math. Soc. 49 (2012), 347-375.
- [DL93] R. DeVore and G.G. Lorentz, Constructive Approximation, Springer Grundlehren, Volume 303, 1993.
- [DiP83] DiPerna, R., Convergence of the viscosity method for isentropic gas dynamics, Commun. Math. Phys. 91 (1983), 1-30.
- [Ey94] G. L. Eyink, Energy dissipation without viscosity in ideal hydrodynamics, I. Fourier analysis and local energy transfer, Phys. D, 78 (1994), 222-240.
- [GHT94] J. Goodman, T. Hou and E. Tadmor, On the stability of the unsmoothed Fourier method for hyperbolic equations Numerische Mathematik 67(1) (1994), 93-129.
- [GO77] D. Gottlieb and S. Orzag, "Numerical Analysis of Spectral Methods : Theory and Applications", SIAM, Philadelphia, PA, 1977.
- [GP03] J.-L Guermond and S. Prudhomme, Mathematical analysis of a spectral hyperviscosity LES model for the simulation of turbulent flows Math. Modeling Numerical Anal. 37 (2003) 893-908.
- [HL07] T. Y. Hou and R. Li, Computing Nearly Singular Solutions Using Pseudo-Spectral Methods, J. Comput. Phys., 226, 379-397, 2007.
- [KK00] G.-S. Karamanos and G. E. Karniadakis, A spectral vanishing viscosity method for large eddysimulations J. of Comput. Physics 163 (2000) 22-50.
- [Kerr93] R.M. Kerr, Evidence for a singularity of the three dimensional, incompressible Euler equations, Phys. Fluids 5 (7) (1993) 1725-1746.
- [Kerr05] R.M. Kerr, Velocity and scaling of collapsing Euler vortices, Phys. Fluids 17 (2005), 075103-114.
- [KH89] R.M. Kerr and F. Hussain, Simulation of vortex reconnection, Physica D 37 (1989) 474.
- [KO72] H.-O Kreiss and J. Oliger, Comparison of accurate methods for the integration of hyperbolic equations. Tellus 24 (1972), 199–215.
- [Kru70] S.N. Kružkov, First order quasilinear equations in several independent variables, Math. USSR Sbornik 10 (1970), 217–243.
- [La86] P. D. Lax, On dispersive difference schemes, Physica D, 18 (1986) 250-254.
- [Lions96] P.L. Lions, "Mathematical Topics in Fluid Mechanics", Volume 1 Incompressible Models, Oxford Lecture Series in Mathematics and its Applications, Oxford (1996).
- [LNT00] M. Lopes Filho, H. J. Nussenzveig and E. Tadmor, Approximate solutions of the incompressible Euler equations with no concentrations Annales De L'institut Henri Poincare (c) Non Linear Analysis, 17 (2000) 371-412.
- [MMO78] Majda, A., McDonough, J. M. and Osher, S., The Fourier Method for Non-Smooth Initial Data, Math. Comp. 32, 1041, 1978.

- [ML78] M. Mock and P. D. Lax, The computation of discontinuous solutions of linear hyperbolic equations, Communications on Pure and Applied Mathematics 31 (1978), 423-430.
- [Mu78] F. Murat, Compacit'e par compensation, Ann. Sc. Norm. Super. Pisa, Cl. Sci., IV. Ser. 5 (1978), 489-507.
- [vN63] von Neumann J, Proposed and analysis of a new numerical method in the treatment of hydrodynamical shock problem, Vol. VI, Collected Works, pp. 361-379, 1963, Pergamon: London.
- [OHFS10] J. Ohlsson, P. Schlatter, P. F. Fischer & D. S. Henningson, Stabilization of the spectral-element method in turbulent flow simulation, ICOSAHOM 2010.
- [ON49] L. Onsager, Statistical hydrodynamics, Nuovo Cimento (9), 6 Supplemento, 2(Convegno Internazionale di Meccanica Statistica) (1949), 279-287.
- [Or72] S. Orszag, Comparison of pseudospectral and spectral approximations, Stud. Appl. Math. 51, (1972), 253-259.
- [Pan94] E. Panov, Uniqueness of the solution of the Cauchy problem for a first-order quasilinear equation with an admissible strictly convex entropy. Mat. Zametki 55 (1994), no. 5, 116–129 (translation in Math. Notes 55 (1994), no. 5-6, 517–525),
- [PSSBS07] R. Pasquetti, E. Severac, E. Serre, P. Bontoux and M. Schafer, From stratified wakes to rotorstator flows by an SVV-LES method Theoretical & Commput. Fluid Dynamics 22(3-4) (2007) 261-273.
- [PNFS13] R. M. Pereira, R. Nguyen van yen, M. Farge and K. Schneider, Wavelet methods to eliminate resonances in the Galerkin-truncated Burgers and Euler equations, Phys. Rev. E 87, 033017, (2013).
- [RFNM11] S. S. Ray, U. Frisch, S. Nazarenko, and T. Matsumuto, Resonance phenomenon for the Galerkintruncated Burgers and Euler equations, Phys. Rev. E 84, 016301 (2011)
- [Sc93] V. Scheffer, An inviscid flow with compact support in space-time, J. of Geom. Anal., 3 (1993), 343–401.
- [SS07] E. Severac and E. Serre, A spectral vanishing viscosity for the LES of turbulent flows within rotating cavities J. Comput. Physics 226 (2007) 1234-1255.
- [Sh97] A. Shnirelman, On the nonuniqueness of weak solutions of the Euler equations, Comm. Pure Appl. Math., 50 (1997), 1261–1286.
- [SK04] S. Sirisup and G. E. Karniadakis, A spectral viscosity method for correcting the long-term behavior of POD models J. of Comput. Physics 194 (2004) 92-116.
- [SW95] C.-W Shu and P.S. Wong, A note on the accuracy of spectral method applied to nonlinear conservation laws J. Scientific Computation 10(3) (1995) 357-369.
- [Tad83] E. Tadmor, The unconditional instability of inflow-dependent boundary conditions in difference approximations to hyperbolic systems, Mathematics of Computation 41 (1983), 309-319.
- [Tad87] E. Tadmor, Stability analysis of finite-difference, pseudospectral and Fourier-Galerkin approximations for time-dependent problems, SIAM Review 29 (1987), 525-555.
- [Tad89] E. Tadmor, Convergence of spectral methods for nonlinear conservation laws, SINUM 26, 1989, pp. 30-44.
- [Tad93a] E. Tadmor, Total variation and error estimates for spectral viscosity approximations Math. Comp. 60 (1993) 245-256.
- [Tad93b] E. Tadmor, "Super viscosity and spectral approximations of nonlinear conservation laws", in Numerical Methods for Fluid Dynamics IV, Proceedings of the 1992 Conference on Numerical Methods for Fluid Dynamics, (M. J. Baines and K. W. Morton, eds.), Clarendon Press, Oxford, 1993, pp. 69-82.
- [Tad94] E. Tadmor, "Spectral methods for hyperbolic problems", Methods numeriques d'ordre eleve pour les ondes en regime transitoire, Lecture notes delivered at Ecole des Ondes, INRIA - Rocquencourt, January 24-28 (1994).
- [Tad02] E. Tadmor, High resolution methods for time dependent problems with piecewise smooth solutions, International Congress of Mathematicians, Proc. of ICM02 Beijing 2002 (Li Tatsien, ed.), Vol.III: Invited Lectures, Higher Education Press, 2002, pp. 747-757.
- [Tad04] E. Tadmor, Burgers' equation with vanishing hyper-viscosity. Commun. Math. Sci. 2(2), 317-324 (2004).
- [Tad07] E. Tadmor, Filters, mollifiers and the computation of the Gibbs phenomenon, Acta Numerica 16 (2007) 305-378.
- [Tar79] L. Tartar, Compensated compactness and applications to partial differential equations, in Research Notes in Mathematics 39, Nonlinear Analysis and Mechanics, Heriott-Watt Symposium, Vol. 4 (R.J. Knopps, ed.)Pittman Press, pp. 136-211 (1979).

[Tar87] L. Tartar, "The compensated compactness method for a scalar hyperbolic equation", Carnegie Mellon Univ. Lecture notes, 87-20, 1987.

(Claude Bardos) UNIVERSITY OF PARIS 7- DENIS DIDEROT LABORATORY JACQUES LOUIS LIONS, UNIVERSITY OF PARIS 6 PARIS, FRANCE *E-mail address*: claude.bardos@gmail.com

(Eitan Tadmor) DEPARTMENT OF MATHEMATICS CENTER OF SCIENTIFIC COMPUTATION AND MATHEMATICAL MODELING (CSCAMM) INSTITUTE FOR PHYSICAL SCIENCES AND TECHNOLOGY (IPST) UNIVERSITY OF MARYLAND MD 20742-4015, USA *E-mail address*: tadmor@cscamm.umd.edu