Computer Physics Communications 207 (2016) 123–135

Contents lists available at ScienceDirect

Computer Physics Communications

journal homepage: www.elsevier.com/locate/cpc

Finite grid instability and spectral fidelity of the electrostatic Particle-In-Cell algorithm

C.-K. Huang [a,*], Y. Zeng [a], Y. Wang [a], M.D. Meyers [a,b,1], S. Yi [a], B.J. Albright [a]

[a] Los Alamos National Laboratory, Los Alamos, NM, 87545, USA

[b] University of California, Los Angeles, CA, 90095, USA

ARTICLE INFO

Article history:
Received 1 September 2015
Received in revised form
7 April 2016
Accepted 23 May 2016
Available online 7 June 2016

Keywords:
Finite grid instability
Particle-In-cell
Spectral fidelity
Numerical instability

ABSTRACT

The origin of the Finite Grid Instability (FGI) is studied by resolving the dynamics in the 1D electrostatic Particle-In-Cell (PIC) model in the spectral domain at the single particle level and at the collective motion level. The spectral fidelity of the PIC model is contrasted with the underlying physical system or the gridless model. The systematic spectral phase and amplitude errors from the charge deposition and field interpolation are quantified for common particle shapes used in the PIC models. It is shown through such analysis and in simulations that the lack of spectral fidelity relative to the physical system due to the existence of aliased spatial modes is the major cause of the FGI in the PIC model.

Published by Elsevier B.V.

1. Introduction

N-body type problems arise in many disciplines and underpins our understanding of complex dynamical systems like plasmas and the cosmos. In a typical N-body problem, the interaction in particle pairs can be of electrostatic, or electromagnetic, or gravitational in nature and each particle responds to a force that is the linear superposition of all one-to-one interactions it receives. Direct calculation of all one-to-one interactions of N_p particles has a computation cost of $O(N_p^2)$, therefore it is amenable to numerical simulation only when N_p is small. The PIC method [1], or more generally the particle-mesh method, is an efficient numerical method that reduces the computation complexity by introducing a computation grid and taking advantage of the linearity in the sum of one-to-one interactions. In the PIC method, the interaction among the particles is mediated by the grid through the Green's function of the interaction represented on the grid. The computation complexity is reduced from $O(N_p^2)$ to $O(N_g)+O(N_p)$, where N_g is the number of grid points. When the number of particles per cell $N_p/N_g \gg 1$, the gain in speedup is large ($\sim N_p$), therefore the PIC method is a popular choice in the *ab-initio* numerical simulation of N-body systems. However, two major problems arise in the PIC method due to the discrete grid: (1) the use of an Eulerian grid for the moments of the particle distribution and fields, in conjunction with individual Lagrangian particles in continuous phase space, implies an inherent inconsistency; (2) the grid representation of the Green's function is usually an approximation of the real Green's function in the continuous space.

Despite the computation efficiency of the PIC method and its wide-spread use, especially in plasma physics, common PIC models are vulnerable to an electrostatic numerical instability known as the Finite Grid Instability (FGI) [2,3] (there is also an electromagnetic instability known as the numerical Cherenkov instability [4–6]). Early practitioners using the electrostatic PIC model to simulate plasma dynamics observed a heating effect to the plasma which depends on the numerical parameters, i.e., the grid size Δx and the number of particles per cell N_C. This numerical heating has been extensively studied since the early development of the PIC model and the empirical scaling of the heating time τ_H of FGI in a thermal plasma, which has the form $\omega_p\tau_H \sim (\lambda_D/\Delta x)^2(N_D+N_C)$, is summarized in Ref. [1]. Here N_D is the number of particles in a Debye length λ_D, ω_p is the plasma frequency. It is also known that FGI comes from the aliased modes in the system due to the incompatibility between the Fourier spectra of the discrete Eulerian and continuous Lagrangian variables. (Numerical Cherenkov instability may also come from an Eulerian–Lagrangian mismatch in the convective derivative [4].) The numerical instabilities have been conventionally analyzed as unphysical resonances between physical and aliased modes [1,4–8]. The locations and growth rates of the

* Corresponding author.
E-mail address: huangck@lanl.gov (C.-K. Huang).

[1] Deceased.

http://dx.doi.org/10.1016/j.cpc.2016.05.021
0010-4655/Published by Elsevier B.V.

unstable modes have been solved for using linear dispersion analysis in limited, yet essential cases, i.e., for spatially uniform cold or Maxwellian distributions. It is worth noting that, unlike the two-stream type instabilities, for some commonly used electrostatic and electromagnetic PIC models, FGI can arise without the intersection of the physical and aliased modes [8,9]. Analysis for more realistic and nonlinear simulations has not been carried out.

Various methods and numerical schemes to mitigate FGI have been proposed, including introducing grid interlacing [10] and random jiggling [10,11], employing higher order particle shapes [9] and temporal/spatial filtering [12], implicit time differencing and enforcing the energy conservation property of the numerical algorithm between time steps [13–16]. All these techniques have shown great promise with regard to the reduction of the instability growth rate. However, this is often achieved with substantial distortion/damping of the meaningful dynamics at the short wavelength scale or by sacrificing conservation properties (such as the loss of momentum conservation in an energy conserving algorithm, which has long been debated in the development of the PIC models [1]). Recent rigorous work on energy conserving algorithm has led to a large improvement of the momentum conservation through nonlinear iterations [14].

The above efforts notwithstanding, the important questions about how and where FGI arises exactly remain to be answered. Previous works treat aliased modes as inherent in the system and study their properties and corresponding mitigation method. However, the origin of the numerical instabilities is clearly unphysical. Therefore, in principle a simulation plasma should be contrasted with the underlying continuous system, which has the same number of particles and particle shape as the simulation plasma, to determine the origin of the numerical instabilities. We call latter the physical system in the following, as it obeys a Vlasov equation for the shaped particles, as long as the same particle shape is used in defining the charge density from the particle and the electric field on the particle.

There are many choices about what to be contrasted between the PIC and the physical systems. Conservation properties and dispersion relationship have been used. It should be noted that one direct consequence of FGI in a PIC model without built-in energy (momentum) conservation property is, as can be expected, the gross violation of the energy (momentum) conservation. For this reason, recent efforts have been devoted to improve the energy and/or momentum conservation of the numerical scheme in order to control FGI. But it should be emphasized that conservation laws are desirable when eliminating the FGI, but they are neither necessary nor sufficient conditions. An isolated plasma system can exhibit various kinds of physical instabilities while strictly conserving momentum and energy. Furthermore, as total energy is a global property of the underlying microscopic processes and only one constraint on the degrees of freedom (two if total momentum is also considered) in the phase space, to understand what gives rise to FGI and its consequences, we need a better resolution into the dynamics. Linear dispersion, in which the eigenmodes with complex Fourier frequency can be viewed as a way to resolve the (linearized) dynamics, is a better choice. However, such analysis is limited to special cases of the particle distribution and small perturbation amplitude. Insight from such analysis for the improvement to the numerical scheme is useful but not easy to obtain and apply to more general situations. Recently, symplectic PIC codes [17,18] have also been developed, for which the symplectic structures of the Hamiltonian system are preserved. The symplectic structure may be a good choice to contrast the PIC and the physical systems, however, it is not clear how it is related to the numerical instabilities at present.

In this paper, we will study the FGI in the 1D electrostatic PIC models by spectrally resolving the dynamics at the single particle level, thus allowing us to identify the components in the model that lead to unphysical instability. The dynamics in PIC result from the superposition of the pair-wise interactions as in a physical system. The major components of an electrostatic PIC model — the charge deposition, the field interpolation and the particle pusher, all operate on a single particle, while the field solver can be viewed as operating on the spatial Fourier modes. Therefore the use of the particle and spectral resolutions are natural choices for this purpose. Such a representation of the PIC models is given in Section 2 and the spectral errors in PIC models are analyzed in Section 3. As an alternative to the individual particle representation, one can also choose the modes of the collective particle motion as a representation. This has the advantage that the plasma dynamics can then be viewed as the collective wave–particle interactions and such couplings in a physical system and in a PIC model can also be contrasted. We note that the deposition, field interpolation and field solver only involve spatial operations at a fixed time, while the particle update in the pusher is a temporal operation in continuous space whose stability and convergence can be verified to rule out its role in FGI. To facilitate simulation comparison with the physical system, the gridless model [19–21] is used, in which all components and elementary operations of the physical system are projected onto the finite Fourier basis. It is demonstrated in Section 4 that the lack of spectral fidelity in the deposition, field interpolation is the major cause of the FGI. Finally we summarize in Section 5.

2. Spectral representation of the PIC model

2.1. Charge deposition

Let us first look at the charge deposition scheme in Fourier space to understand the effect of aliasing. We will see that the most important effect in Fourier space is the summation over all Brillouin zones which is the result of a convolution process between a continuous spectrum and a periodic spectrum over Brillouin zones. The sampling needed to go from continuous space to a discrete grid is the cause of the latter spectrum.

In a grid-based model like PIC, the contribution of a particle at position $\boldsymbol{x_p}$[2] and of total charge Q on the density grid $\boldsymbol{r_\rho}$ is $\rho(\boldsymbol{r_\rho}) = QW(\boldsymbol{r_\rho}, \boldsymbol{x_p})$, where $\boldsymbol{r_\rho}$ is a vector on a uniform grid with grid size $\Delta\boldsymbol{x} = (\Delta x, \Delta y, \Delta z)$. The interpolation function for the deposition is $W(\boldsymbol{r_\rho}, \boldsymbol{r}) = W(|\boldsymbol{r_\rho} - \boldsymbol{r}|)$. Note that we have assumed the cell volume $V = \Delta x\Delta y\Delta z = 1$ and dropped it for clarity. The difference between a particle shape function $S(\boldsymbol{r})$ and an interpolation function $W(\boldsymbol{r_\rho}, \boldsymbol{r})$ is discussed in Appendix A. In the rest of this section, we will not distinguish these two and will use $S(\boldsymbol{r})$ for clarity, $\rho(\boldsymbol{r_\rho}) = QS_\rho(\boldsymbol{r_\rho} - \boldsymbol{x_p})$. We define a transform (note this is not necessarily the proper Discrete Fourier Transform as $\boldsymbol{r_\rho}$ may be shifted from the origin of the coordinate system, as will be shown later),

$$\tilde{\rho}(\boldsymbol{k}) = \sum_{\boldsymbol{r_\rho}} \rho(\boldsymbol{r_\rho}) e^{-i\boldsymbol{k}\cdot\boldsymbol{r_\rho}} = Q \sum_{\boldsymbol{r_\rho}} S_\rho(\boldsymbol{r_\rho} - \boldsymbol{x_p}) e^{-i\boldsymbol{k}\cdot\boldsymbol{r_\rho}}, \quad (2.1)$$

which can be viewed as the continuous Fourier transform of $\rho(\boldsymbol{r}) \sum_{\boldsymbol{r_\rho}} \delta(\boldsymbol{r} - \boldsymbol{r_\rho})$, where $\rho(\boldsymbol{r}) = QS_\rho(\boldsymbol{r} - \boldsymbol{x_p})$. Then the non-unitary inverse transform is

$$\rho(\boldsymbol{r_\rho}) = \frac{1}{8\pi^3} \int_{-\infty}^{\infty} \tilde{\rho}(\boldsymbol{k}) e^{i\boldsymbol{k}\cdot\boldsymbol{r_\rho}} d\boldsymbol{k}.$$

[2] Italic bold font is used for vectors.