Chapter 1

Numerical Methods Project: The Boris Push in PIC Codes

1.1 The Relevant Equations

Particle-in-cell codes are nowadays the most popular popular tool for simulating plasma systems. One of the best references for what they are, how they work and how reliable the results are is the book by Birdsall and Langdon 1995, which describes the main numerical methods used and some of their properties. However, Particle-in-cell codes have evolved greatly in the last two decades and new techniques and optimizations have been produced and even put in practice. Even so, with the exception of quasistatic codes, they are still involved in solving the same physical equations. As such, it is useful for anyone interested in working with such software to know and understand the principles behind.

In general, PIC codes have four main componets:

1. A Maxwell solver which propagates the Maxwell equations (which are relativistic invariant by themselves) in time and space on the grid

$$\boldsymbol{\nabla} \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0} \tag{1.1a}$$

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0 \tag{1.1b}$$

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{1.1c}$$

$$\boldsymbol{\nabla} \times \mathbf{B} = \mu_0 \mathbf{j} + \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t}\,; \tag{1.1d}$$

2. A field gatherer that interpolates the electromagnetic field at the particle positions on the grid;

3. A particle pusher that advances the positions and velocities of each particle under the action of the Lorentz force (which we will write in its relativistic form)

$$\frac{\mathrm{d}\mathbf{p}_\alpha}{\mathrm{d}t} = q_\alpha \left(\mathbf{E} + \frac{\mathbf{p}_\alpha}{m_\alpha \gamma_\alpha} \times \mathbf{B} \right) \tag{1.2a}$$

$$\gamma_\alpha = \sqrt{1 + \left(\frac{\mathbf{p}_\alpha}{m_\alpha c} \right)^2}\,, \tag{1.2b}$$

where α indexes each particle;

4. Acurrent and charge depositor which computes the current and charge densities on the grid by interpolating the particle distributions.

The main appeal of this approach is its self-consistency. That is, the total fields used are both those that are part of the electromagnetic waves that are introduced in the system (in general laser beams) and those generated by the charged particles that compose the plasma. As such we also include the long range Coulombian interaction between particles. The short range interaction, namely collisions between particles, is by default neglected since we usually simulate rarefied plasmas, but many codes now come with additional routines that include these processes. Additional routines are now developed with the advent of the high intensity laser technologies because at the corresponding energies reached by the particles quantum electrodynamical effects become relevant. Although there are quite a few PIC codes that include QED routines, there is still a long way untill these algorithms reach the efficiency and stability that of those four main ones described above. As such, the implementation of QED effects in numerical plasma simulations is currently a hot research topic.

It is mandatory to mention that while the four steps above outline a microscopic model, PIC simulations are not completely microscopic due to technological limitations regarding computing power. Instead of working with one virtual particle for one real particle, it is customary to use macro-particles. A macro-particle represents many particles of the same species (from 10^6 to 10^11 depending on the propeties of our plasma) moving colectively. These particles are obviously not localized at a single point, but rather they have a shape function attached to them to make the derivation of currents and charge densities more consistent. For a long time the use of macro-particles was not supported by argument and was a source of criticism towards PIC methods. The defense was built only on the excuse of that the simulations give very accurate statistical results. Things are different nowadays. We can now explain (quite easily in fact) that the macro-particles themselves can be interpreted a statistical ensamble of real particles. The secret lies in the Vlasov equation.

1.1.1 The Connection with the Vlasov Equation

Let us revisit the Vlasov equation, which we derived in **??**

$$\frac{\partial \rho}{\partial t} + \mathbf{f} \cdot \frac{\partial \rho}{\partial \mathbf{p}} + \mathbf{v} \cdot \frac{\partial \rho}{\partial \mathbf{r}} = 0 \,, \tag{1.3}$$

where ρ was the distribution function that describes the entire system of particles, $\mathbf{p} = (\mathbf{p_1}, \mathbf{p_2}, \dots)$ and $\mathbf{v} = (\mathbf{v_1}, \mathbf{v_2}, \dots)$, $\mathbf{r} = (\mathbf{r_1}, \mathbf{r_2}, \dots)$ the momenta, the velocities, and the positions of the particles, and $\mathbf{f} = (\mathbf{f_1}, \mathbf{f_2}, \dots)$ the forces acting on each particle.

The main argument in the following discussion is a relativistic upgrade of the one in Liu, Tripathi, and Eliasson 2020. For consistency with the equations we outlined for the PIC method, I rewrite this equation in its relativistic form and considering all forces to be of the Lorentz type

$$\frac{\partial \rho}{\partial t} + \sum_{\alpha} \left[q_\alpha \left(\mathbf{E}_\alpha + \frac{\mathbf{p}_\alpha}{m_\alpha \gamma_\alpha} \times \mathbf{B}_\alpha \right) \frac{\partial \rho}{\partial \mathbf{p}_\alpha} + \frac{\mathbf{p}_\alpha}{m_\alpha \gamma_\alpha} \frac{\partial \rho}{\partial \mathbf{r}_\alpha} \right] = 0 \,, \tag{1.4}$$

where α indexes all the particles in the system and the fields E_α and B_α are to be computed at the position of particle α.

The key insight now is that imposing that the particles move under the Newton-Lorentz equation (1.2a) implies having a stationary solution for the distribution function. That is, the equations

$$\frac{\mathrm{d}\mathbf{p}_\alpha}{\mathrm{d}t} = q_\alpha \left(\mathbf{E} + \frac{\mathbf{p}_\alpha}{m_\alpha \gamma_\alpha} \times \mathbf{B} \right) \tag{1.5a}$$

$$\frac{\mathrm{d}\mathbf{r}_\alpha}{\mathrm{d}t} = \frac{\mathbf{p}_\alpha}{m_\alpha \gamma_\alpha}\,, \tag{1.5b}$$

reduce the Vlasov equation as follows

$$\frac{\partial \rho}{\partial t} + \sum_\alpha \left(\frac{\partial \rho}{\partial \mathbf{p}_\alpha} \frac{\mathrm{d}\mathbf{p}_\alpha}{\mathrm{d}t} + \frac{\partial \rho}{\partial \mathbf{r}_\alpha} \frac{\mathrm{d}\mathbf{r}_\alpha}{\mathrm{d}t} \right) = \frac{\mathrm{d}\rho}{\mathrm{d}t} = 0\,. \tag{1.6}$$

Of course this is not an equivalency. While equation (1.5) implies that the distribution function of the entire system is stationary, the reverse is not true, unless we do a rough approximation and suppose that the total distribution can be separated in a sum of independent single particle distribution functions. By employing this latter approximation we would unavoidably neglect some intrinsic interactions that take place in our system. Nonetheless, this problem does not affect the validity of our Particle-in-cell method. The thing is that while equation (1.5) doesn't describe all the complete stationary solution of the Vlasov equation, it still describes at least one particular stationary solution. Working with superparticles is like studying the evolution of an ensable of these solutions. Thus, by including a relevant (yet not large enough to give unreasonable simulation times) number of superparticles we obtain a statistically realistic solution. Some even call PIC a Monte-Carlo method because of this.

The reverse approach is used in numerical studies of plasma physics using Vlasov codes (VC). These approaches simply try to solve the Vlasov equation as it is in order to obtain exact solutions (exact up to numerical errors). If one has the total distribution function then every statistical piece of information about the system is known. Some basics about how this can be achieved and computational optimization can be found in Silin 2020. However, directly solving such a big solution with a number of variables proportional to the real number of particles in the system takes a lot of time if we try to simulate realistically sized systems.

An approach based on the splitting of the distribution function is not completely flawed. One can try to get closer to reality by using a better approximation by expanding the total distribution in a series that contains single-particle terms, two-particle terms, and so on. While it is true that this improves the solutions greatly, the cost in computation time is equally great and much refinement would have to be done.

The presentation so far should not give the reader the impression that PIC is the supperior approach. In fact, numerical heating is a common problem of PIC codes and the stability conditions for the simulations are quite restrictive, which together with the computational limitations reduce the amount of experiments you can run. An easy to follow rought sketch of the trade off between PIC and VC and a discussion on when is one better than the other is found in Bertrand *et al.* 2005.

1.2 Methods used in Particle-in-cell simulations

1.2.1 The Boris Push

It is time now to tackle the first method in the context of Particle-in-cell simulations. It is discussed in Birdsall and Langdon 1995, there are different ways to implement particles pushers, either by partially separating the action of the electric and magmetic fields (Buneman 1967), or by separating them completely (Boris 1970). Along the years, even more alternatives have

surfaced. A nice up-to-date review of the currently relevant relativistic schemes for particle dynamics in electromagnetic fields is Ripperda *et al.* 2018. In this section we will present and analyze the Boris push, since it is widely implemented in the currently available PIC codes. The popularity of this specific scheme stems from its practical performance. Since its initial proposal, the Boris push was observed to have good accuracy over long integration times in simulations. The explanation for why it works so well was given only recently in Qin *et al.* 2013. They studied only the properties of the non-relativistic scheme. Because of this, we will also start with the classical case, wich is simpler, in order to gain some intuition.

Classical Boris Push

We aim to solve the followwing equations

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \mathbf{v} \tag{1.7a}$$

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \frac{q}{m}\left(\mathbf{E} + \mathbf{v}\times\mathbf{B}\right). \tag{1.7b}$$

We will employ the Störmer-Verlet method

$$\mathbf{r_{n+1}} = \mathbf{r_n} + \mathbf{v}_{\mathbf{n+\frac{1}{2}}}\Delta t \tag{1.8a}$$

$$\mathbf{v}_{\mathbf{n+\frac{1}{2}}} = \mathbf{v_n} + \frac{q}{m}\frac{\Delta t}{2}\left[\mathbf{E} + \mathbf{v}\times\mathbf{B}\right]_{\mathbf{r=r_n}} \tag{1.8b}$$

$$\mathbf{v_n} = \mathbf{v}_{\mathbf{n+\frac{1}{2}}} + \frac{q}{m}\frac{\Delta t}{2}\left[\mathbf{E} + \mathbf{v}\times\mathbf{B}\right]_{\mathbf{r=r_{n+1}}}. \tag{1.8c}$$

In order to find a good evaluation for the quantities enclosed in square brackets we need to play around a bit with these equations. By lowering the third one by one step and replacing in the second we get

$$\frac{\mathbf{v}_{\mathbf{n+\frac{1}{2}}} - \mathbf{v}_{\mathbf{n-\frac{1}{2}}}}{\Delta t} = \frac{q}{m}\left[\mathbf{E} + \mathbf{v}\times\mathbf{B}\right]_{\mathbf{r=r_n}}. \tag{1.9}$$

Let us now take equation (1.8a) at two consecutive steps

$$\mathbf{r_n} = \mathbf{r_{n-1}} + \mathbf{v}_{\mathbf{n-\frac{1}{2}}}\Delta t \tag{1.10a}$$

$$\mathbf{r_{n+1}} = \mathbf{r_n} + \mathbf{v}_{\mathbf{n+\frac{1}{2}}}\Delta t \tag{1.10b}$$

and add them to obtain

$$\frac{\mathbf{r_{n+1}} - \mathbf{r_{n-1}}}{\Delta} = \frac{\mathbf{v}_{\mathbf{n+\frac{1}{2}}} + \mathbf{v}_{\mathbf{n-\frac{1}{2}}}}{2}, \tag{1.11}$$

which is we will use as an approximation for the velocity at $\mathbf{r} = \mathbf{r_n}$. We finally reach the following scheme

$$\frac{\mathbf{r_{n+1}} - \mathbf{r_n}}{\Delta t} = \mathbf{v}_{\mathbf{n+\frac{1}{2}}} \tag{1.12a}$$

$$\frac{\mathbf{v}_{\mathbf{n+\frac{1}{2}}} - \mathbf{v}_{\mathbf{n-\frac{1}{2}}}}{\Delta t} = \frac{q}{m}\left(\mathbf{E_n} + \frac{\mathbf{v}_{\mathbf{n+\frac{1}{2}}} + \mathbf{v}_{\mathbf{n-\frac{1}{2}}}}{2}\times\mathbf{B_n}\right). \tag{1.12b}$$

This is the Boris push, but it is not yet the final algorithm. In order to be able to implement this scheme in an actual code, we need a way to separate $\mathbf{v}_{\mathbf{n}+\frac{\mathbf{1}}{\mathbf{2}}}$ and $\mathbf{v}_{\mathbf{n}-\frac{\mathbf{1}}{\mathbf{2}}}$ from the second equation. There are many equivalent ways to do so, but Boris originally came up with a method to separate the action of the electric and magnetic fields by employing the following three equations

$$\mathbf{v}^- = \mathbf{v}_{\mathbf{n}-\frac{\mathbf{1}}{\mathbf{2}}} + \frac{q}{m}\mathbf{E}_\mathbf{n}\frac{\Delta t}{2} \tag{1.13a}$$

$$\frac{\mathbf{v}^+ - \mathbf{v}^-}{\Delta t} = \frac{q}{m}\frac{\mathbf{v}^+ + \mathbf{v}^-}{2} \times \mathbf{B}_\mathbf{n} \tag{1.13b}$$

$$\mathbf{v}_{\mathbf{n}+\frac{\mathbf{1}}{\mathbf{2}}} = \mathbf{v}^+ + \frac{q}{m}\mathbf{E}_\mathbf{n}\frac{\Delta t}{2}\,, \tag{1.13c}$$

where the second step rewritten in the following way:

$$\mathbf{v}^+ = \mathbf{v}^- + (\mathbf{v}^- + \mathbf{v}^- \times \mathbf{t}) \times \mathbf{s} \tag{1.14a}$$

$$\mathbf{t} = \frac{q}{2m}\mathbf{B}_\mathbf{n}\Delta t \tag{1.14b}$$

$$\mathbf{s} = \frac{2}{1+t^2}\mathbf{t}\,. \tag{1.14c}$$

The full classical Boris push code recipe is described by equations (1.12a), (1.13a), (1.13c) and (1.14) put together. Note that we basically work only with mid-step velocities. But we can always recover the actual velocities from equation (1.8b) like this

$$\mathbf{v}_\mathbf{n} = \mathbf{v}_{\mathbf{n}+\frac{\mathbf{1}}{\mathbf{2}}} - \frac{q}{m}\frac{\Delta t}{2}\left(\mathbf{E}_\mathbf{n} + \frac{\mathbf{v}_{\mathbf{n}+\frac{\mathbf{1}}{\mathbf{2}}} + \mathbf{v}_{\mathbf{n}-\frac{\mathbf{1}}{\mathbf{2}}}}{2} \times \mathbf{B}_\mathbf{n}\right)\,. \tag{1.15}$$

Such an equation is useful in order to get the initial conditions for the mid-step velocities from the initial conditions given for the velocity.

We introduced equations (1.13) and (1.14) out of nowhere, but it is quite straightforward to see that by eliminating v^- and v^+ we simply recover equation (1.12b). This is just a convenient way to write for a machine to understand, but is simply a reorganization of the equation (1.12b). It does not influence numerical properties. That is, all the nunerical properties of the algorithm lie in equation (1.12). There is a geometrical interpretation to the break up described by equation (1.13). v^- and v^+ give the drift motion due to electric field, equation (1.13) described a rotation under the effect of a constant magnetic field. For a more in depth description with images to help visualization I recommend the Master Thesis of Micluță-Câmpeanu 2019.

Relativistic Boris Push

The relativistic version should discretize equation (1.2a) in the same way we presented so far. If we make the notation $\mathbf{u} = \gamma \mathbf{v}$, the equations of the algorithm are now simply

$$\frac{\mathbf{r_{n+1}} - \mathbf{r_n}}{\Delta t} = \frac{\mathbf{u_{n+\frac{1}{2}}}}{\gamma_{n+\frac{1}{2}}},\ \gamma_{n+\frac{1}{2}} = \sqrt{1 + \left(\frac{u_{n+\frac{1}{2}}}{c}\right)^2} \tag{1.16a}$$

$$\mathbf{u}^- = \mathbf{u_{n-\frac{1}{2}}} + \frac{q}{m}\mathbf{E_n}\frac{\Delta t}{2} \tag{1.16b}$$

$$\mathbf{u}^+ = \mathbf{u}^- + (\mathbf{u}^- + \mathbf{u}^- \times \mathbf{t}) \times \mathbf{s},\ \mathbf{t} = \frac{q}{2m\gamma_n}\mathbf{B_n}\Delta t,\ \mathbf{s} = \frac{2}{1+t^2}\mathbf{t} \tag{1.16c}$$

$$\gamma_n = \sqrt{1 + \left(\frac{\mathbf{u}^-}{c}\right)^2} = \sqrt{1 + \left(\frac{\mathbf{u}^+}{c}\right)^2} \tag{1.16d}$$

$$\mathbf{u_{n+\frac{1}{2}}} = \mathbf{u}^+ + \frac{q}{m}\mathbf{E_n}\frac{\Delta t}{2}\,. \tag{1.16e}$$

1.2.2 Symplecticity and Volume Conservation Theory

In terms of accuracy, the Boris push remains at its core a twist on the Störmer-Verlet method, so it an be shown that it is a second order scheme. But there are things that are not inherited this way. We already saw that the Störmer-Verlet and leapfrog integrators are very similar. It a known fact that the leapfrog scheme is one of the simplest symplectic and as such it has outstanding conservation properties. But we can not expect the Boris push to be the same. The thing that throws everything off is the dependence on velocity that we have introduced in the expression of the force. Even so, the Boris push still has a strength in that it is volume preserving. It is also important to mention that it conserves energy exactly in the absence of the electric field. This may not seem that great in itself, since the electric field is never vanishing in practical simulations, but it was a hint towards the idea that it might be volume preserving. In what follows we will delve a bit into the concept of symplecticity and see how we can find out if a scheme has this property. Of course, our example for this theory will be the Boris push algorithm.

In order to give a concrete mathematical definition for simplecticity we have to define a handy tool first, as presented on p. 164 in Arnold 1997.

Definition 1. Let M^{2n}, $n \in \mathbb{N}$, be a differentiable manifold with an even number of dimensions (this is general, so we can use any such manifold, but for our numerical methods related endeavours we really only need to talk about $\mathbb{R}^{2n}$). **An exterior form of degree two** (or a **2-form**) on this manifold is a map $\omega_2 : M^n \times M^n \to M$ that is bilinear and skew symmetric:

$$\omega_2(\lambda_1 \boldsymbol{\xi_1} + \lambda_2 \boldsymbol{\xi_2}, \boldsymbol{\xi_3}) = \lambda_1 \omega_2(\boldsymbol{\xi_1}, \boldsymbol{\xi_3}) + \lambda_2 \omega_2(\boldsymbol{\xi_2}, \boldsymbol{\xi_3})$$
$$\omega_2(\boldsymbol{\xi_1}, \boldsymbol{\xi_2}) = -\omega_2(\boldsymbol{\xi_2}, \boldsymbol{\xi_1})\,,$$

$\forall \lambda_1,\ \lambda_2 \in M, \boldsymbol{\xi_1},\ \boldsymbol{\xi_2},\ \boldsymbol{\xi_3} \in M^n$. If we also have the extra property that $\omega_2(\boldsymbol{\xi_1}, \boldsymbol{\xi_2}) = 0,\ \forall \boldsymbol{\xi_2}$ implies $\boldsymbol{\xi_1} = 0$ we say that the 2-form is non-degenerate.

Note that the even dimensionality of our manifold is an important aspect of the definition. It is also important to remark that all the concepts we discuss here in this chapter play an important role in analytical mechanics, since the phase space is always even dimensional.

In a two dimensional space, the determinant of the matrix obtained by the augmentation of two vectors gives the area of the paralelogram described by the two vectors (Golomb 1985). The determinant in this context is also a basic example of a 2-form. This is not a mere coincidence. In general, a differential 2-form at a point (an exterior 2-form on the tangent space at that point) computes a local oriented differential area there. This observation is in a way valid for any k-form, but we are not interested in developing too much geometry in this text.

I mentioned the tangent space. This concept is actually necessary to understand in order to move forward, so let us present a short definition adapted to our interests (Weisstein 2020a):

Definition 2. Let $\boldsymbol{x}$ be a point in our manifold M^{2n}. If we attach a copy of $\mathbb{R}^{2n}$ tangential to M^{2n} at $\mathbf{x}$ we obtain a structure called **the tangent space** of M^{2n} at $\boldsymbol{x}$ and we denote it by $T_{\boldsymbol{x}}M$.

As any concept in differential geometry, it is easy to understand it in spaces with a small number of dimensions. The simplest to visualize in my opinions is the 2D surface of a sphere. If we choose any point on the sphere and stick at that point an infinite plane we obtain the tangent space at that point. The idea of the tangent space is something that people are actually used with from calculus. Say we have a curve on our manifold that passes through $\mathbf{x}$. The derivative of the curve at $\mathbf{x}$ is a vector in the tangent space. Coming back to our 2D example, on a shpere surface the derivative of a curve on it at a point is tangent to the sphere, so basically an stright arrow tangent to the sphere at that point. But clearly that arrow is not *on* the sphere. It is on the tangent plane part of the tangent space. This is really an extension to the idea in calculus that the derrivative gives the slope of the function at a point.

Definition 3. A symplectic form on M^{2n} is a smooth closed non-degenerate 2-form ω_2 on M^{2n} such that the alternating bilinear map $\omega_2^{\boldsymbol{x}} : T_{\boldsymbol{x}}M^{2n} \times T_{\boldsymbol{x}}M^{2n} \to \mathbb{R}$ defined by the expression of ω_2 at every point $\boldsymbol{x} \in M^{2n}$ is also non-degenerate.

This may seem quite an abstract and hard to digest idea, so let us dismiss it with a particularization relevant for our problem at hand. If our manifold is $\mathbb{R}^{2n}$ then any 2-form $\omega_2 : \mathbb{R}^{2n} \times \mathbb{R}^{2n} \to \mathbb{R}$ is symplectic simply if $\omega_2(\mathbf{x}, \mathbf{x}) = 0$ for any $\mathbf{x}$ in $\mathbb{R}^n$ (Weisstein 2020b).

Definition 4. A linear map $A : M^n \to M^n$ is called *symplectic* if there exists a symplectic form $\omega_2 : M^n \times M^n \to M$ such that $\omega_2(A\boldsymbol{\xi}, A\boldsymbol{\eta}) = \omega_2(\boldsymbol{\xi}, \boldsymbol{\eta})\,,\ \forall \boldsymbol{\xi}, \boldsymbol{\eta} \in M^n\,.$

For real manifolds we can reformulate this in the following way:

Definition. A linear map $A : \mathbb{R}^n \to \mathbb{R}^n$ is called *symplectic* if there exists a 2-form ω_2 on $\mathbb{R}^{2n}$ with $\omega_2(\boldsymbol{x}, \boldsymbol{x}) = 0,\ \forall \boldsymbol{x} \in \mathbb{R}^n$ such that $\omega_2(A\boldsymbol{\xi}, A\boldsymbol{\eta}) = \omega_2(\boldsymbol{\xi}, \boldsymbol{\eta})\,,\ \forall \boldsymbol{\xi}, \boldsymbol{\eta} \in \mathbb{R}^n\,.$

An equivalent condition can be given in matrix form

$$A^T J A = J\,, \tag{1.17}$$

where $J = \begin{bmatrix} 0_n & I_n \\ -I_n & 0_n \end{bmatrix}$ and I_n is the n-dimensional unitary matrix (remark: $J^{-1} = -J$).

Taking the determinant of each side in the identity 1.17 results that the determinant of the matrix A is ± 1. We can do better. It is always 1. The proof usually involves using the Pfaffian, but it is possible to avoid this (Rim 2018). Let us make use of the fact that $A^T A$ is symmetric (note that $(A^T A)^T = A^T (A^T)^T = A^T A$). Since $\det(A) \neq 0$, A is invertible, so $A^T A$ is also positive definite.

As a side note we should mention that the simple fact that we found that A is invertible lets us compute the inverse. Indeed,

$$A^T J A = J \Rightarrow A^T J A A^{-1} = A^T J = J A^{-1} \Rightarrow A^{-1} = J^{-1} A^T J \,. \tag{1.18}$$

We will do a little trick now. $A^T A$ is both symmetric and positive definite. This means that its eigenvalues are real and positive. If $\mathbf{v}$ is an eigenvector of $A^T A$ and λ the corresponding eigenvalue. Then $(A^T A + I_{2n})\mathbf{v} = (\lambda + 1)\mathbf{v}$. This means that the eigenvalues of $A^T A + I_{2n}$ are greater than one. Since the determinant is the product of eigenvalues, we have

$$\det\big(A^T A + I_{2n}\big) > 1 \,.$$

We can extract an A^T now

$$A^T A + I_n = A^T (A + (A^T)^{-1}) = A^T (A + J^{-1} A J) \,,$$

such that

$$0 < 1 < \det(A) \det\big(A + J^{-1} A J\big) \,. \tag{1.19}$$

Let us write our matrix as $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, with $a,\ b,\ c,\ d \in \mathbb{R}^n$. Now

$$A + J^{-1} A J = \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} 0_n & -I_n \\ I_n & 0_n \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0_n & I_n \\ -I_n & 0_n \end{bmatrix} =$$

$$= \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} 0_n & -I_n \\ I_n & 0_n \end{bmatrix} \begin{bmatrix} -b & a \\ -d & c \end{bmatrix} =$$

$$= \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} d & -c \\ -b & a \end{bmatrix} = \begin{bmatrix} a+d & b-c \\ -(b-c) & a+d \end{bmatrix} .$$

Denoting $B \equiv a + d$ and $C \equiv b - c$ we can make use of the following decomposition

$$A + J^{-1} A J = \begin{bmatrix} \frac{1}{\sqrt{2}} I_n & \frac{1}{\sqrt{2}} I_n \\ \frac{\mathrm{i}}{\sqrt{2}} I_n & -\frac{\mathrm{i}}{\sqrt{2}} I_n \end{bmatrix} \begin{bmatrix} B + \mathrm{i}C & 0_n \\ 0_n & B - \mathrm{i}C \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} I_n & \frac{1}{\sqrt{2}} I_n \\ \frac{\mathrm{i}}{\sqrt{2}} I_n & -\frac{\mathrm{i}}{\sqrt{2}} I_n \end{bmatrix}^T ,$$

Which we can use in equation (1.19)

$$0 < \det(A) \det(B + \mathrm{i}C) \det(B - \mathrm{i}C) = \det(A) |\det(B + \mathrm{i}C)| \,,$$

which concludes that $\det(A) > 0$.

Together with the fact $\det(A)$ can only take the values ± 1 this completes the proof that **the determinant of the associated matrix of a symplectic linear map is 1**.

This is the complete set of mathematical definitions and nomenclature that develop the concept of symplectic maps. But a question still remains: what is a symplectic map *really*? Well, it is a generalization of area-preservance (the image of a subset of our manifold through an area-preserving map has the same volume as the subset itself; in this thesis we use volume and area preservence interchangably). In practice it is useful to use the following result in order to decide if this property holds: **a linear map is area-preserving if and only if its associated matrix has the determinant equal to 1**. The proof is omitted since it can be easily found in many places across the internet (also, note that area-preservance is a necessary, but not sufficient condition for symplecticity, considering our previous discussion). Now, looking at phase space, if a linear map is symplectic then, informally, the sum of areas projected on the planes $(\mathbf{q_i}, \mathbf{p_i})$, $i = \overline{1, n}$ is conserved (Weisstein 2020c), so we can have a lot more quatities conserved. Symplecticity, in this perspective, could be related to generalized angular momenta

conservation. Indeed, it has been shown that there are symplectic algorithms (Störmer-Verlet schemese, or modifications of it) which conserve angular momentum exactly (M.-Q. Zhang and Skeel 1995). We must mention that energy conservation is better than symplecticity. In general energy conservation sits above symplecticity, which sits above area-preservance. This may not seem that obvious, since all these three conservation properties are identical in two dimensions and the 2D phase space (1D system) is pretty much the only one our mind can really visualize (a 2D system would have a 4D phase space). The advantage of symplectic or at least volume preserving algorithms is that the energy drift is not diverging and is quite small.

We still have to make one more step in order to study the properties of the Boris push. The problem is that the map that describes this scheme is not linear, so we can not directly use what we presented so far. We simply have to define simplecticity for a larger class of mappings (p. 183 Hairer, Lubich, and Wanner 2006).

Definition 5. A map $f : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is symplectic if its Jacobian matrix is symplectic.

In a similar manner we can define a area-preservance for a wider selection of maps.

Definition 6. A map $f : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is area-preserving if the determinant of its Jacobian matrix is equal to 1.

No matter how we look at all these things, what actually do every time is to check the condition 1.17 for a matrix. We usually call the matrices that solve this identity symplectic. It turns out that for a fixed n the set of all the symplectic matrices form a group (the identity element is I_{2n}, The inverse is given by equation (1.18), proving closure and associativity is trivial). In fact, it is a Lie group commonly denoted as $Sp(2n, \mathbb{R})$. The tangent space at the identity element defines its Lie algebra.

1.2.3 Conservation Properties of the Boris push

In order to check the conservation properties of the Boris push we need to find a way to express equation (1.12) as a matrix equation between the position and velocity at a step and those at the next step. For this we will use the hat map (while it is connected to group theory, we introduce it just as a tool).

Definition 7. For a vector $\boldsymbol{v} = (v_1, v_2, v_3) \in \mathbb{R}^3$ the corresponding hat map is the matrix

$$\hat{\boldsymbol{v}} = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix} .$$

Proposition 1. The vector product of two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ can be expressed with the help of the hat map in the following way

$$\boldsymbol{u} \times \boldsymbol{v} = \hat{\boldsymbol{u}}\boldsymbol{v} .$$

With this preparation we can extract from the Boris push equations

$$\frac{\mathbf{r_{n+1}} - \mathbf{r_n}}{\Delta t} = \mathbf{v_{n+\frac{1}{2}}} \tag{1.20a}$$

$$\frac{\mathbf{v_{n+\frac{1}{2}}} - \mathbf{v_{n-\frac{1}{2}}}}{\Delta t} = \frac{q}{m}\left(\mathbf{E_n} + \frac{\mathbf{v_{n+\frac{1}{2}}} + \mathbf{v_{n-\frac{1}{2}}}}{2} \times \mathbf{B_n}\right) \tag{1.20b}$$

the step update into a more convenient matrix notation

$$\mathbf{r_{n+1}} = \mathbf{r_n} + \Delta t \mathbf{v}_{\mathbf{n+\frac{1}{2}}} \tag{1.21a}$$

$$\left(I_3 + \frac{1}{2}\frac{q\Delta t}{m}\hat{\mathbf{B}}_\mathbf{n}\right)\mathbf{v}_{\mathbf{n+\frac{1}{2}}} = \left(I_3 - \frac{1}{2}\frac{q\Delta t}{m}\hat{\mathbf{B}}_\mathbf{n}\right)\mathbf{v}_{\mathbf{n-\frac{1}{2}}} + \frac{q\Delta t}{m}\mathbf{E_n}\,, \tag{1.21b}$$

where I_3 is the three dimensional identity matrix and

$$\hat{\mathbf{B}}_\mathbf{n} = \begin{bmatrix} 0 & -B_n^3 & B_n^2 \\ B_n^3 & 0 & -B_n^1 \\ -B_n^2 & B_n^1 & 0 \end{bmatrix}.$$

Since the determinant of $\left(I_3 + \frac{1}{2}\frac{q\Delta t}{m}\hat{\mathbf{B}}_\mathbf{n}\right)$ is simply $1+\left(\frac{1}{2}\frac{q\Delta t}{m}B_n^1\right)^2+\left(\frac{1}{2}\frac{q\Delta t}{m}B_n^2\right)^2+\left(\frac{1}{2}\frac{q\Delta t}{m}B_n^3\right)^2 \neq 0$ it has an inverse, so we can further reduce equation (1.21b) to

$$\mathbf{v}_{\mathbf{n+\frac{1}{2}}} = \left(I_3 + \frac{1}{2}\frac{q\Delta t}{m}\hat{\mathbf{B}}_\mathbf{n}\right)^{-1}\left(I_3 - \frac{1}{2}\frac{q\Delta t}{m}\hat{\mathbf{B}}_\mathbf{n}\right)\mathbf{v}_{\mathbf{n-\frac{1}{2}}} + \frac{q\Delta t}{m}\left(I_3 + \frac{1}{2}\frac{q\Delta t}{m}\hat{\mathbf{B}}_\mathbf{n}\right)^{-1}\mathbf{E_n}\,. \tag{1.22}$$

For convenience we will make the notation

$$R \equiv \left(I_3 + \frac{1}{2}\frac{q\Delta t}{m}\hat{\mathbf{B}}_\mathbf{n}\right)^{-1}\left(I_3 - \frac{1}{2}\frac{q\Delta t}{m}\hat{\mathbf{B}}_\mathbf{n}\right), \tag{1.23}$$

so the system of equations for our scheme is now

$$\mathbf{r_{n+1}} = \mathbf{r_n} + \Delta t R\mathbf{v}_{\mathbf{n-\frac{1}{2}}} + \frac{q\Delta t^2}{m}\left(I_3 + \frac{1}{2}\frac{q\Delta t}{m}\hat{\mathbf{B}}_\mathbf{n}\right)^{-1}\mathbf{E_n} \tag{1.24a}$$

$$\mathbf{v}_{\mathbf{n+\frac{1}{2}}} = R\mathbf{v}_{\mathbf{n-\frac{1}{2}}} + \frac{q\Delta t}{m}\left(I_3 + \frac{1}{2}\frac{q\Delta t}{m}\hat{\mathbf{B}}_\mathbf{n}\right)^{-1}\mathbf{E_n}\,. \tag{1.24b}$$

The Jacobi matrix of this scheme is

$$\frac{\partial(\mathbf{r_{n+1}}, \mathbf{v}_{\mathbf{n+\frac{1}{2}}})}{\partial(\mathbf{r_n}, \mathbf{v}_{\mathbf{n-\frac{1}{2}}})} = \begin{bmatrix} \frac{\partial \mathbf{r_{n+1}}}{\partial \mathbf{r_n}} & \frac{\partial \mathbf{r_{n+1}}}{\partial v_{n-\frac{1}{2}}} \\ \frac{\partial \mathbf{v}_{\mathbf{n+\frac{1}{2}}}}{\partial \mathbf{r_n}} & \frac{\partial \mathbf{v}_{\mathbf{n+\frac{1}{2}}}}{\partial v_{n-\frac{1}{2}}} \end{bmatrix} = \begin{bmatrix} I_3 + \Delta t \frac{\partial \mathbf{v}_{\mathbf{n+\frac{1}{2}}}}{\partial \mathbf{r_n}} & R\Delta t \\ \frac{\partial \mathbf{v}_{\mathbf{n+\frac{1}{2}}}}{\partial \mathbf{r_n}} & R \end{bmatrix}. \tag{1.25}$$

We now have to check if

$$\left[\frac{\partial(\mathbf{r_{n+1}}, \mathbf{v}_{\mathbf{n+\frac{1}{2}}})}{\partial(\mathbf{r_n}, \mathbf{v}_{\mathbf{n-\frac{1}{2}}})}\right]^T J \frac{\partial(\mathbf{r_{n+1}}, \mathbf{v}_{\mathbf{n+\frac{1}{2}}})}{\partial(\mathbf{r_n}, \mathbf{v}_{\mathbf{n-\frac{1}{2}}})} = J\,. \tag{1.26}$$

Under the notation

$$\frac{\partial(\mathbf{r_{n+1}}, \mathbf{v}_{\mathbf{n+\frac{1}{2}}})}{\partial(\mathbf{r_n}, \mathbf{v}_{\mathbf{n-\frac{1}{2}}})} = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}, \tag{1.27}$$

this identity can be expressed as

$$\left[\frac{\partial(\mathbf{r_{n+1}},\mathbf{v}_{\mathbf{n+\frac{1}{2}}})}{\partial(\mathbf{r_n},\mathbf{v}_{\mathbf{n-\frac{1}{2}}})}\right]^T \begin{bmatrix} 0 & I_3 \\ -I_3 & 0 \end{bmatrix} \frac{\partial(\mathbf{r_{n+1}},\mathbf{v}_{\mathbf{n+\frac{1}{2}}})}{\partial(\mathbf{r_n},\mathbf{v}_{\mathbf{n-\frac{1}{2}}})} = \begin{bmatrix} A_1^T & A_3^T \\ A_2^T & A_4^T \end{bmatrix} \begin{bmatrix} 0 & I_3 \\ -I_3 & 0 \end{bmatrix} \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} =$$

$$= \begin{bmatrix} A_1^T & A_3^T \\ A_2^T & A_4^T \end{bmatrix} \begin{bmatrix} A_3 & A_4 \\ -A_1 & -A_2 \end{bmatrix} = \begin{bmatrix} A_1^T A_3 - A_3^T A_1 & A_1^T A_4 - A_3^T A_2 \\ A_2^T A_3 - A_4^T A_1 & A_2^T A_4 - A_4^T A_2 \end{bmatrix} = \begin{bmatrix} 0 & I_3 \\ -I_3 & 0 \end{bmatrix},$$

which leads to the following system of matrix equations

$$\left(A_1^T A_3\right)^T = A_1^T A_3 \tag{1.28a}$$

$$\left(A_2^T A_4\right)^T = A_2^T A_4 \tag{1.28b}$$

$$\left(A_1^T A_4 - A_3^T A_2\right)^T = A_1^T A_4 - A_3^T A_2 = I_3\,. \tag{1.28c}$$

It is quite impossible to work with these equations, since they can be written for any electromagnetic fields. However, if we want to show that the Boris push is not symplectic, a particular example that doesn't satisfy these identities is enough. Indeed, if we make the choice that the electric and magnetic fields do not depend on position, then $\frac{\partial \mathbf{v}_{\mathbf{n+\frac{1}{2}}}}{\partial \mathbf{r_n}} = 0$, so the A matrices become

$$A_1 = I_3 \tag{1.29a}$$

$$A_2 = R\Delta t \tag{1.29b}$$

$$A_3 = 0 \tag{1.29c}$$

$$A_4 = R\,. \tag{1.29d}$$

With this choice, equation (1.28c) leads to the obvious contradiction

$$A_1^T A_4 - A_3^T A_2 = R = I_3\,, \tag{1.30}$$

so the Boris push is not symplectic. But is it area-preserving? To see this we only have to compute the determinant of the Jacobian (1.25) and see if it is equal to 1.

$$\left|\frac{\partial(\mathbf{r_{n+1}},\mathbf{v}_{\mathbf{n+\frac{1}{2}}})}{\partial(\mathbf{r_n},\mathbf{v}_{\mathbf{n-\frac{1}{2}}})}\right| = \begin{vmatrix} I_3 + \Delta t \frac{\partial \mathbf{v}_{\mathbf{n+\frac{1}{2}}}}{\partial \mathbf{r_n}} & R\Delta t \\ \frac{\partial \mathbf{v}_{\mathbf{n+\frac{1}{2}}}}{\partial \mathbf{r_n}} & R \end{vmatrix} \overset{row1-\Delta t\, row2}{=\!=} \begin{vmatrix} I_3 & 0_3 \\ \frac{\partial \mathbf{v}_{\mathbf{n+\frac{1}{2}}}}{\partial \mathbf{r_n}} & R \end{vmatrix} = |R|\,. \tag{1.31}$$

Let us denote $\frac{1}{2}\frac{q\Delta t}{m}\hat{\mathbf{B}}_\mathbf{n}$ by Ω. By direct calculation we can show that

$$\det(I_3+\Omega) = \det(I_3-\Omega) = 1 + \left(\frac{1}{2}\frac{q\Delta t}{m}B_n^1\right)^2 + \left(\frac{1}{2}\frac{q\Delta t}{m}B_n^2\right)^2 + \left(\frac{1}{2}\frac{q\Delta t}{m}B_n^3\right)^2, \tag{1.32}$$

such that

$$|R| = \det\big((I_3+\Omega)^{-1}(I_3-\Omega)\big) = \det\big((I_3+\Omega)^{-1}\big)\det(I_3-\Omega) = \frac{\det(I_3-\Omega)}{\det(I_3+\Omega)} = 1\,. \tag{1.33}$$

This concludes the proof that the Boris push is area-preserving.

We will not bother to study also the relativistic algorithm since it is pretty much the same thing with a few changes. In the article by R. Zhang *et al.* 2015, the authors claim to provide a *new* particle pusher that is volume preserving (since, as they say, the relativistic Boris push is not), but it turns out that their original scheme is exactly that of Boris. Nonetheless, the article proves area-preservance for the relativistic Boris push.