

Numerical simulation of beam dynamics in particle accelerators



A flavor of techniques and tools used to predict/push the performance of the Large Hadron Collier at CERN **Part 1: single particle methods**

Giovanni Iadarola and Riccardo De Maria, Beams Department, CERN

CERN Prévessii

ATL



- Particle accelerators
 - Examples of applications
 - Working principles

• Beam optics calculations

- An example: the LHC betatron squeeze
- Numerical optimization techniques

• Particle tracking

- Motivations
- Need for symplectic methods
- $\circ~$ Experience with GPU computing



A **particle accelerator** is a machine used to **accelerate particles** (electrons, protons, ions, etc...) using **electromagnetic fields**

→ There are many accelerators in the word used for a variety of applications (industrial, medical, research)

Example of industrial application:

• **Ion implantation** in the fabrication of integrated circuits







Example of medical application:

• Accelerator for cancer treatment





Example of research application:

• Large Hadron Collider (LHC) for high-energy physics research

Introduction

Largest and most powerful particle accelerator ever built (27 km circumference) It is designed to:

- store **450 GeV protons** (in two counter rotating beams)
- accelerate them up to 7 TeV
- collide the two beams in four points of the ring (for high energy physics experiments)







The Large Hadron Collider (LHC) 8-fold symmetric structure: 8 Long Straight Sections (LSS) to host experiments and other equipment 8 Arcs (2.45 km each - Periodic magnet)



CÉRN





- Particle accelerators
 - o Examples of applications
 - Working principles
- Beam optics calculations
 - An example: the LHC betatron squeeze
 - Numerical optimization techniques
- Particle tracking
 - Motivations
 - Need for symplectic methods
 - Experience with GPU computing

- CERN
- To manipulate and accelerate charged particles we use electromagnetic fields

• The Lorentz force acts on the particles \mathbf{F} =

 $\mathbf{F} = q\mathbf{E} + q\mathbf{v} imes \mathbf{B}$

- The magnetic field force does not change the energy of the particles
- The acceleration itself needs to be done by an electric field in specifically designed accelerating structures



Accelerating structures can be concatenated to form a **linear accelerator** ("linac"):

- → Acceleration is **very fast** (single passage)
- → But achievable energy is quite limited





- To manipulate and accelerate charged particles we use electromagnetic fields
 - \circ The Lorentz force acts on the particles $\mathbf{F} = q\mathbf{E} + q\mathbf{v} imes \mathbf{B}$

 Magnetic fields do not change the energy of the particles but can be used very effectively to guide them





• "Dipole magnets" are used to bend the particles' trajectory



• We can use a set of **dipole magnets** to keep the particles on a **closed trajectory**





- We can use a set of **dipole magnets** to keep the particles on a **closed trajectory**
 - Allows to accumulate "re-use" for many turns the energy gain from the same accelerating structure (RF cavity)
- In the Large Hadron Collider (LHC):
 - We want to increase the proton energy by ~6000 GeV
 - Accelerating cavities provide only ~500 keV/turn (in average)
 - Acceleration is done in about
 ~15million turns (20 minutes)





•

Due to the way they are produced, particle beams always have a small divergence





- Due to the way they are produced, particle beams always have a small divergence
 - Over time particles would be inevitably lost on the walls of the beam-pipe







- Due to the way they are produced, particle beams always have a small divergence
 - Over time particles would be inevitably lost on the walls of the beam-pipe
- To keep the particles close to the center of the beam-pipe we use quadrupole magnets
 - In a quadrupole, the magnetic force is linearly proportional to the distance from the center of the magnet ($F_x = -kx$)





- Due to the way they are produced, particle beams always have a small divergence
 - Over time particles would be inevitably lost on the walls of the beam-pipe
- To keep the particles close to the center of the beam-pipe we use quadrupole magnets
 - In a quadrupole, the magnetic force is linearly proportional to the distance from the center of the magnet ($F_x = -k x$) → it acts like a focusing lens
 - Quadrupole magnets focus the beam in one direction but defocus on the orthogonal one





- Due to the way they are produced, particle beams always have a small divergence
 - Over time particles would be inevitably lost on the walls of the beam-pipe
- To keep the particles close to the center of the beam-pipe we use quadrupole magnets
 - In a quadrupole, the magnetic force is linearly proportional to the distance from the center of the magnet ($F_x = -k x$) → it acts like a focusing lens
 - Quadrupole magnets focus the beam in one direction but defocus on the orthogonal one → quadrupoles with opposite polarity are alternated to confine the beams in both planes







- Particle accelerators
 - Examples of applications
 - Working principles
- Beam optics calculations
 - An example: the LHC betatron squeeze
 - Numerical optimization techniques
- Particle tracking
 - Motivations
 - Need for symplectic methods
 - Experience with GPU computing



- In the presence of **dipole and quadrupole magnets alone** (linear regime) it is possible to **compute the envelope of the beam** without having to evaluate the trajectories of the single particles
 - \circ $\,$ Courant-Snyder formalism based on the Floquet theorem

$$x(s) = \sqrt{\varepsilon} \sqrt{\beta(s)} \cos[\Psi(s) + \phi]$$

• These calculations are called in general "linear optics" calculations





Quadrupole strengths can be used to shape the particle beam envelope (in the same way in which lenses can be used to shape a beam of light)

Example: LHC betatron squeeze





Quadrupole strengths can be used to shape the particle beam envelope (in the same way in which lenses can be used to shape a beam of light)

Example: LHC betatron squeeze





- Particle accelerators
 - Examples of applications
 - Working principles

• Beam optics calculations

- An example: the LHC betatron squeeze
- Numerical optimization techniques
- Particle tracking
 - Motivations
 - Need for symplectic methods
 - Experience with GPU computing





- The relation between the quadrupole strengths and the beam size is non-trivial
 - Several technical constraints also need to be taken into account (e.g. maximum quadrupole strength, current sign, size of the beam pipes, magnets in series that must have the same current)
 - **Numerical optimizers** need to be used to identify suitable quadrupole strengths as a function of given constraints on the beam envelope

CERN's workhorse code for these calculations is the **MAD-X code**.





- To illustrate how an optimization algorithm works we consider a **simple problem** with **2 constraints** and **2 degrees of freedom**:
 - We want the maximum and the minimum of the beam envelope $\sigma(s)$ to assume specified values σ_A and σ_B (marked by the red lines in figure)
 - We can change the strength of **two families of quadrupoles** (k_{q1} and k_{q2})



We define a suitable "cost function":

$$F\left(k_{\mathrm{Q1}},k_{\mathrm{Q2}}\right) =$$

$$\sqrt{\left(\frac{\sigma_{\max}\left(k_{\mathrm{Q1}},k_{\mathrm{Q2}}\right)-\sigma_{\mathrm{A}}}{\sigma_{\mathrm{A}}}\right)^{2}+\left(\frac{\sigma_{\min}\left(k_{\mathrm{Q1}},k_{\mathrm{Q2}}\right)-\sigma_{\mathrm{B}}}{\sigma_{\mathrm{B}}}\right)^{2}}$$

To solve our problem we need to search the minimum of this quantity as function of k_{Q1} and k_{Q2}

This is called an "optimization problem"



We define a suitable "cost function":

$$F(k_{\rm Q1}, k_{\rm Q2}) = \sqrt{\left(\frac{\sigma_{\rm max}(k_{\rm Q1}, k_{\rm Q2}) - \sigma_{\rm A}}{\sigma_{\rm A}}\right)^2 + \left(\frac{\sigma_{\rm min}(k_{\rm Q1}, k_{\rm Q2}) - \sigma_{\rm B}}{\sigma_{\rm B}}\right)^2}$$

To solve our problem we need to search the minimum of this quantity as function of $k_{\rm Q1}$ and $k_{\rm Q2}$

- With only two degrees of freedom we can visualize the function as surface
- With more than two degrees of freedom it can become too expensive to map the whole parameter space → we need to search for the minimum "blindly"









CERN

1. At the given point we evaluate the gradient of the cost function, $-\nabla F = -\left(\frac{\partial F}{\partial k_{Q1}}, \frac{\partial F}{\partial k_{Q2}}\right)$ tells us the direction in which our surface is the steepest



- 1. At the given point we evaluate the gradient of the cost function, $-\nabla F = -\left(\frac{\partial F}{\partial k_{Q1}}, \frac{\partial F'}{\partial k_{Q2}}\right)$ tells us the direction in which our surface is the steepest
- 2. We take a **new point in that direction** and we go back to 1.

CERN



- 1. At the given point we evaluate the gradient of the cost function, $-\nabla F = -\left(\frac{\partial F}{\partial k_{Q1}}, \frac{\partial F'}{\partial k_{Q2}}\right)$ tells us the direction in which our surface is the steepest
- 2. We take a **new point in that direction** and we go back to 1.

CERN



- 1. At the given point we evaluate the gradient of the cost function, $-\nabla F = -\left(\frac{\partial F}{\partial k_{Q1}}, \frac{\partial F'}{\partial k_{Q2}}\right)$ tells us the direction in which our surface is the steepest
- 2. We take a **new point in that direction** and we go back to 1.

CERN

After a certain number of iterations the algorithm will converge to a minimum of the cost function



- 1. At the given point we evaluate the gradient of the cost function, $-\nabla F = -\left(\frac{\partial F}{\partial k_{Q1}}, \frac{\partial F}{\partial k_{Q2}}\right)$ tells us the direction in which our surface is the steepest
- 2. We take a **new point in that direction** and we go back to 1.

CERN

After a certain number of iterations the algorithm will converge to a minimum of the cost function





Similar techniques are user to shape the beam trajectories (closed orbit)





Iterative methods are used also **to correct the beam trajectory** (with respect to a known reference) due to daily small fluctuations → done **online on the circulating beams**

Before correction



After correction



Outline



- Particle accelerators
 - Examples of applications
 - Working principles
- Beam optics calculations
 - An example: the LHC betatron squeeze
 - Numerical optimization techniques

• Particle tracking

- o Motivations
- Need for symplectic methods
- Experience with GPU computing



- Dipolar and quadrupolar fields are in principle sufficient to keep the particles on a closed trajectory and keep them focused.
- Nevertheless in a realistic accelerator the situation is more complex:
 - Magnets are not perfect (dipole and quadrupole magnets have unwanted deviation from the ideal field shapes)
 - Magnets are not "exactly" where they are supposed to be (alignment errors)
 - Particles **do not have all exactly the same energy** (typical relative spread ~10⁻³)
 - Need of "chromatic corrections" using sextupole magnets
- ightarrow A realistic machine has **unavoidable non-linearities**



Sextupole magnet



- In the presence of these effects, the particle motion gets much more complex:
 - The envelope equation is not anymore enough
 - Depending on the initial conditions particles can be lost after a certain number of turns
- We need to **numerically simulate the motion of the particle** in the accelerator:
 - We are interested in quantifying how many particles will be lost over a realistic time → for the LHC we need to simulate ~millions of turns!



Outline



- Particle accelerators
 - Examples of applications
 - Working principles
- Beam optics calculations
 - An example: the LHC betatron squeeze
 - Numerical optimization techniques

• Particle tracking

- Motivations
- Need for symplectic methods
- Experience with GPU computing



- Simulations on such long time scales are subject to particular issues, which we will illustrate using a simple example:
 - We assume uniform focusing force





- Simulations on such long time scales are subject to particular issues, which we will illustrate using a simple example:
 - We assume uniform focusing force
 - In such a field the **particle oscillate** around the axis x = 0



Equations of motion

$$F = ma$$

$$\frac{dv_x}{dt} = -\frac{k}{m}x$$
$$\frac{dx}{dt} = v_x$$



Such a system **preserves the initial energy of the particle**, defined as:

$$E = \frac{1}{2}mv_x^2 + \frac{1}{2}kx^2 = \text{const}$$

Kinetic energy Potential energy

Proof:

$$\frac{dE}{dt} = mv_x \frac{dv_x}{dt} + kx \frac{dx}{dt}$$
$$= mv_x \left(\frac{dv_x}{dt} + \frac{k}{m}x\right) = 0$$



We compare **two numerical methods** to compute x(t):

<u>Method 1</u>: We use a numerical integration method to find an **approximated solution** to the **exact problem**

Eq. of motion in vector form

We introduce a **discrete time-step** Δt

$$\frac{d\mathbf{z}}{dt} = \mathbf{f}(\mathbf{z})$$

with:
$$\mathbf{z}(t) = \begin{pmatrix} x(t) \\ v_x(t) \end{pmatrix}$$

Runge-Kutta scheme:

$$\mathbf{z}_{n+1} = \mathbf{z}_n + \frac{1}{6} \left(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4 \right)$$

with:
$$\mathbf{k}_1 = \Delta t \ \mathbf{f}(\mathbf{z}_n),$$

$$\mathbf{k}_2 = \Delta t \ \mathbf{f} \left(\mathbf{z}_n + \frac{\mathbf{k}_1}{2} \right)$$

$$\mathbf{k}_3 = \Delta t \ \mathbf{f} \left(\mathbf{z}_n + \frac{\mathbf{k}_2}{2} \right)$$

$$\mathbf{k}_4 = \Delta t \ \mathbf{f} \left(\mathbf{z}_n + \mathbf{k}_3 \right)$$



We compare **two numerical methods** to compute x(t):

<u>Method 2</u>: We find an approximated problem for which we are able to compute the exact solution



We concentrate our focusing force at discrete locations ("lenses")







The Runge-Kutta method is more accurate on a short time interval





- The Runge-Kutta method is more accurate on a short time interval
- On very long time-spans ...





- The Runge-Kutta method is more accurate on a short time interval
- On very long time-spans ...





- The Runge-Kutta method is more accurate on a short time interval
- On very long time-spans ...





- The Runge-Kutta method is more accurate on a short time interval
- On very long time-spans ...





- The Runge-Kutta method is more accurate on a short time interval
- On very long time-spans ...





- The Runge-Kutta method is more accurate on a short time interval
- On very long time-spans ...





- The Runge-Kutta method is more accurate on a short time interval
- On very long time-spans ...





- The Runge-Kutta method is more accurate on a short time interval
- On very long time-spans the Runge-Kutta method slowly "consumes" the energy of the particles





- The Runge-Kutta method is more accurate on a short time interval
- On very long time-spans the Runge-Kutta method slowly "consumes" the energy of the particles





- The Runge-Kutta method is more accurate on a short time interval
- On very long time-spans the Runge-Kutta method slowly "consumes" the energy of the particles
 - "Fake" physical phenomena are introduced
 - Runge-Kutta cannot be used to predict slow effects on the beam
- In spite of being less accurate on short times, the symplectic scheme does not suffer from this issues





- In general, for long-term tracking we do **need to use symplectic algorithms**:
 - → The numerical solution needs to preserve fundamental properties of the physical system such as energy conservation



Outline



- Particle accelerators
 - Examples of applications
 - Working principles
- Beam optics calculations
 - An example: the LHC betatron squeeze
 - Numerical optimization techniques

• Particle tracking

- Motivations
- Need for symplectic methods
- \circ $\,$ Experience with GPU computing



In accelerator studies we are interested **tracking a large number of particles** (~100 000) to probe different initial conditions:

- Simulation of each particle is independent
- We are facing an "embarrassingly parallel problem"



→ Particularly suited for **GPU acceleration**



- Graphics Processing Units (GPUs) are chips developed since the 80s to perform graphics calculations (video rendering), typically installed on a video-card
 - Main applications are gaming and computer graphics in general
- Since the early 2000s GPU vendors provide tools to use the GPUs also for **generalpurpose parallel computing:**
 - Libraries and tools to exploit these resources have flourished
 - Cards dedicated to high performance computing have been commercialized





CPU Optimized for Serial Tasks



GPU Accelerator Optimized for Parallel Tasks



- Has a small number of complex computing cores (up to 8)
- Very fast clock rates
- Can access large memory (> 100 GB)

- Has a large number of simpler computing cores (>1000)
- Slower clock rates
- Can access relatively small memory (~16 GB)

https://www.olcf.ornl.gov/wp-content/uploads/2018/06/intro_to_HPC_gpu_computing.pdf



Resources allocation:

- A GPU has more resources dedicated to Arithmetic-Logic operations (ALUs) compared to a covariational CPU
- A GPU has less resources dedicated to control and cache memory

Control	ALU	ALU							
	ALU	ALU							
Cache									
DRAM									

Control	ALU									
Cache										
Control	ALU									
Cache										
Control	ALU									
Cache										
Control	ALU									
Cache										
DRAM										

CPU

GPU

https://www.olcf.ornl.gov/wp-content/uploads/2018/06/intro_to_HPC_gpu_computing.pdf



CPU

optimized for speed (but reduced capacity)

GPU

optimized for capacity
(but reduced speed)



Which is better depends on your needs...

https://www.olcf.ornl.gov/wp-content/uploads/2018/06/intro_to_HPC_gpu_computing.pdf





https://github.com/SixTrack/sixtracklib





https://github.com/SixTrack/sixtracklib





https://github.com/SixTrack/sixtracklib





https://github.com/SixTrack/sixtracklib



When tracking a **large number of particles GPUs become very attractive** (still the price of the device can be more expensive...)







https://github.com/SixTrack/sixtracklib



- A **particle accelerator** uses **electromagnetic field** to accelerate and manipulate charged particles
 - Accelerating structures are used to increase the energy of the particles
 - **Dipole magnets** are used to keep the beams on a closed trajectory
 - **Quadrupole magnets** are used to confine (focus) the particles
- In the presence of **linear forces alone** it is possible to **compute the beam envelope** (optics) without computing the single particle trajectories
 - **Quadrupole strengths** can be **used to shape the particle beam envelope** (in the same way in which lenses can be used to shape a beam of light)
 - Numerical optimizers (like the gradient method) need to be used to identify suitable quadrupole strengths as a function of given constraints on the beam envelope
- **Particle tracking** is the simulation of individual particles in the accelerator over a very large number of turns:
 - **Symplectic algorithms** are required in order to preserve fundamental properties of the physical system
 - **GPU computers** are particularly suited for this kind of simulations



- So far we have studied **"single-particle" methods**, which neglect the interactions among circulating particles
- In the second part we will focus on methods for "collective effects", which are particularly relevant when the beam is very intense (large number of particles)



Thanks for your attention!