# PHY 564
# Advanced Accelerator Physics
# Lecture 25

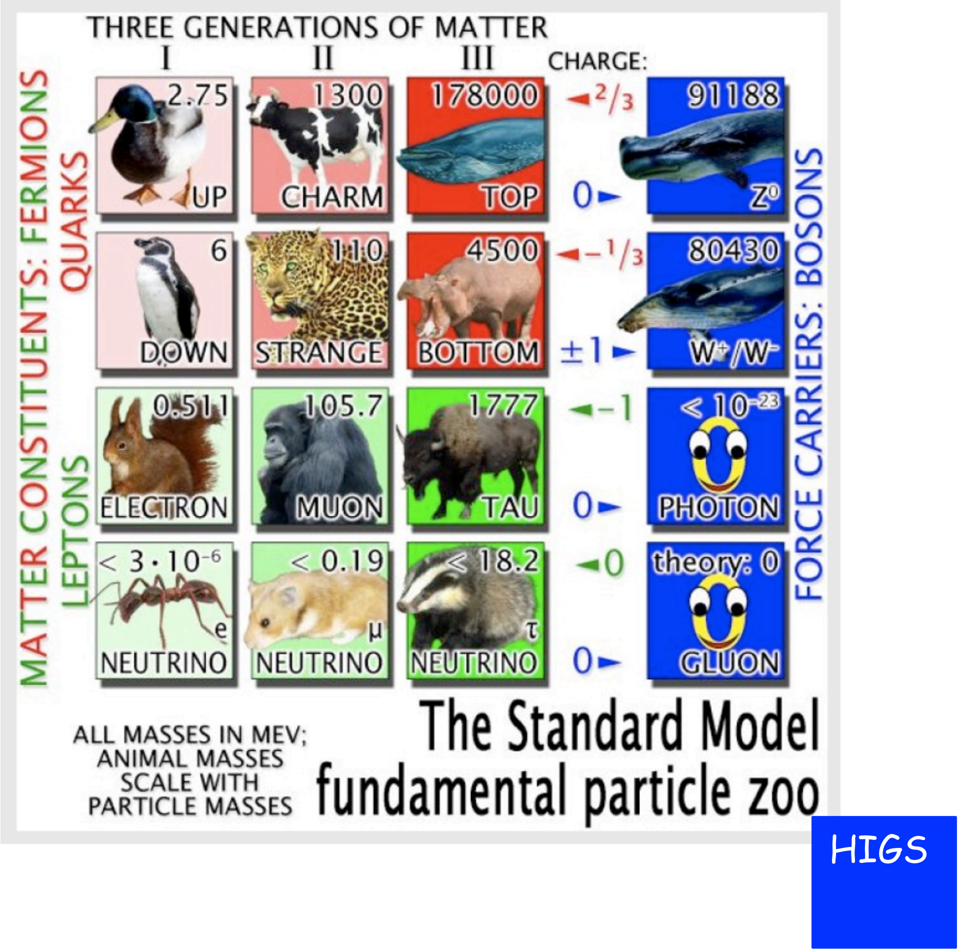
# Colliders and non-linear resonances

Vladimir N. Litvinenko

CENTER for ACCELERATOR SCIENCE AND EDUCATION
Department of Physics & Astronomy, Stony Brook University

**Lecture 26. Colliders**

Colliders are probably the most fundamental application of the accelerators. From first days of accelerators, they were used for collider experiments to discover new particles (from positrons or antiproton to Higgs) and new forces (weak, strong). There are two main parameters of any collider: the center-of-mass energy of collisions (e.g. energy available for generating new particles) and luminosity, which is defining the collider "productivity". Generally speaking, there are two types of collisions used for generating new particles: (a) beam colliding with a fixed targets (e.g. at rest in the lab, e.g. the detector, frame) and (b) beam-beam collisions (when both colliding species are moving with relativistic velocities in the lab frame. Second type of collisions became the leader in high-energy physics, starting from pioneering electron and electron-position colliders built in Novosibirsk and SLAC in late 1960s.



Each and every type of accelerators was used for collision experiments with fixed target – they are still popular and are used for, what is now caller, low and medium energy nuclear/particle physics. Fig. 1 shows detector for 12 GeV electrons coming from CEBAF recirculating linac and colliding with a fixed target in front of the detector.

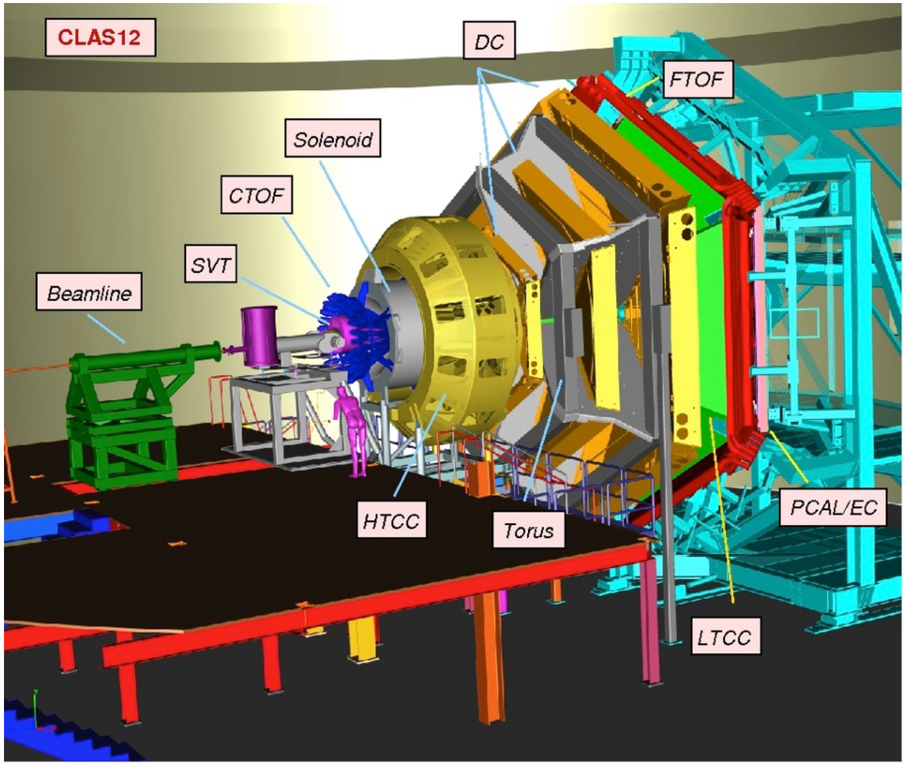



Fig. 1. CLAS12 detector at Hall B at CEBAF, Thomas Jefferson National Accelerator Facility, Newport News, VA

Let's consider kinematics of such collisions, shown in Fig.2.

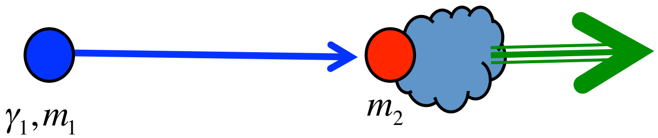



Fig. 2. A relativistic particle collides with a particle at rest. Resulting "products", shown as a cloud, carries the total momentum of the incoming particle – it means that there is always non-zero kinetic energy in the rest frame.

As we learned from relativistic mechanics, the maximum energy available for generating particle is given by the relativistic invariant contraction of 4-momentum:

$$P^i = p_1^i + p_2^i = \left( \frac{E_1 + E_2}{c}, \vec{p}_1 + \vec{p}_2 \right);$$
$$E_{c.m.} \equiv Mc^2 = \sqrt{P^i P_i} = \sqrt{\left(E_1 + E_2\right)^2 - c^2 \left(\vec{p}_1 + \vec{p}_2\right)^2} \tag{1}$$

Note that $\vec{p}_1 + \vec{p}_2 \neq 0$ always reduces the available energy. The meaning of (1) is to indicate the threshold of the available energy. In other words, if resulting products of collision are at rest, the maximum total mass of the products cannot exceed $M$. Hence, if we are looking for a new particle with mass of 1 TeV, we will need at least 1 TeV c.m. energy. If particles are generated in pairs (for example as particle-antiparticle pair) – we will need 2 TeV c.m. In reality, we always need more that the threshold energy.

Now, let look at the kinematic in Fig.2:

$$E_1 = \gamma_1 m_1 c^2; c\vec{p}_1 = \hat{z}\beta_1 E_1; \beta_1 = \sqrt{1-\gamma_1^{-2}}; E_2 = m_2 c^2; \vec{p}_2 = 0; \gamma_2 = 1;$$
$$M^2 = \frac{E_{c.m.}{}^2}{c^4} = \left(\gamma_1 m_1 + m_2\right)^2 - \beta_1^2 \gamma_1^2 m_1^2 = m_1^2 + m_2^2 + 2\gamma_1 m_1 m_2; \tag{2}$$
$$E_{c.m.} = c^2 \sqrt{m_1^2 + m_2^2 + 2\gamma_1 m_1 m_2}.$$

First two terms are just rest energies of the colliding particles and for the most interesting case of highly relativistic collisions $\gamma_1 >> 1$

$$E_{c.m.} \cong c^2 \sqrt{2\gamma_1 m_1 m_2} + O\left(\gamma_1^{-1}\right) \approx \sqrt{2E_1 m_2 c^2} \tag{3}$$

scaling of the available energy in such collisions is very unfavorable: to increase energy available for generating new particles 10-fold one need to increase accelerator energy 100-fold.

Colliding beams with fixed target has tremendous advantage – you can collide a relatively weak beam with solid or gaseous target and take advantage of essentially infinite (just look at Avogadro number of $6\ 10^{23}$ …) number of particles at rest. This is why for a while fixed target were the only collisions used for nuclear and high energy physics. The energy of accelerators was increasing exponentially by invention of new concepts: from electrostatic MeV-scale in 1930s to GeV scale synchrotrons in 1950s. But eventually machines became so large and scaling so unfavorable that invention of circular colliders became a necessity and later, a reality.

The main disadvantage of colliding beams is that we have a very limited number of particles in a single bunch – typically in $10^9$-$10^{11}$ range – and necessity to maintain particles from diffusing and eventually being lost. At the same time, the main advantage of colliders is favorable energy scaling when particles moving towards each other:

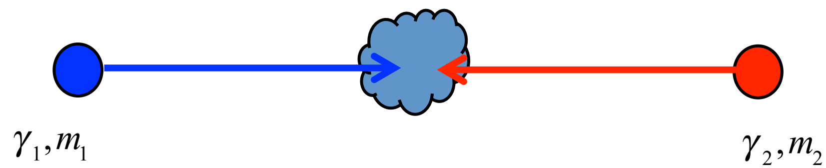



Fig. 3. Head-on collision of two particles.

$$E_1 = \gamma_1 m_1 c^2; c\vec{p}_1 = \hat{z}\beta_1 E_1; \beta_1 = \sqrt{1-\gamma_1^{-2}}; E_2 = \gamma_1 m_2 c^2; \vec{p}_2 = -\hat{z}\beta_2 E_2; \beta_{1,2} = \sqrt{1-\gamma_{1,2}^{-2}};$$

$$\gamma_i^2\left(1-\beta_i^2\right) \equiv 1 \rightarrow E_i^2 - c^2\vec{p}_i^{\,2} \equiv \left(m_i c^2\right)^2;$$

$$M^2 = \frac{E_{c.m.}^{\;\;2}}{c^4} = \left(\gamma_1 m_1 + \gamma_2 m_2\right)^2 - \left(\beta_1\gamma_1 m_1 - \beta_2\gamma_2 m_2\right)^2 = \tag{4}$$

$$m_1^2 + m_2^2 + 2\gamma_1\gamma_2 m_1 m_2\left(1+\beta_1\beta_2\right);$$

$$E_{c.m.} = c^2\sqrt{m_1^2 + m_2^2 + 2\left(1+\beta_1\beta_2\right)\gamma_1\gamma_2 m_1 m_2}.$$

and for the most interesting case of highly relativistic collisions $\gamma_i >> 1,\ 1-\beta_i << 1$

$$E_{c.m.} \cong 2c^2\sqrt{\gamma_1\gamma_2 m_1 m_2} + O\left(\gamma_i^{-1}\right) \approx 2\sqrt{E_1 E_2} \tag{5}$$

the energy scales as a geometric average of the energies of colliding particles. In this case increasing energies of particles 10-fold give 10-fold increase in c.m. energy.

The most obvious case is of colliding particles with the same mass (electron-positions, proton-proton or proton-antiproton) and the same energy:

$$E_1 = E_2 = E = \gamma mc^2; c\vec{p}_1 = -c\vec{p}_2 = \hat{z}\beta E; P^i = \left(2E, \vec{0}\right);$$
$$E_{c.m.} = 2E; \tag{6}$$

e.g. the total energy of colliding particles is available.

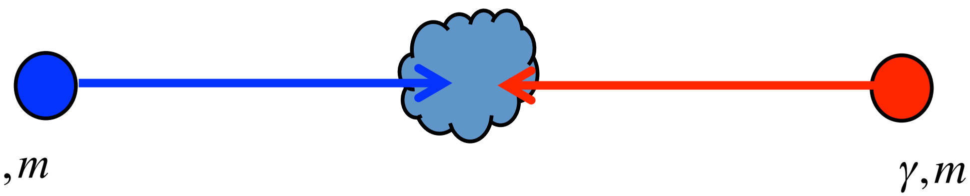



Fig. 4. Head-on collision of two particles with the same masses and energies.

It is especially true for electron-position (or proposes muon-anti-muon, $\mu\bar{\mu}$ ) colliders, where electron and positron can annihilate and make all their energy available for generate other species of particles. It is also true that lepton colliders ($e^-e^+$ or $\mu\bar{\mu}$ ) are the cleanest – we collide indeed elementary particles and start with well-defined initial states (including spin) and well-defined energies. The main limitation of electron-position circular colliders is relatively low energy (max ~ 100 GeV per beam) when compared with pp colliders (max ~ TeV). Muon colliders, while being studies as a potentially important high energy tool, have two main problems:

(a) we cannot generate low emittance muons and have to find a very complex (and untested) cooling techniques to bring the beam quality to acceptable value;

(b) muons are leaving (in their rest-frame) only for 2.2 microseconds – hence, they have to be accelerated and cooled very quickly …

These are indeed very challenging problems.

**Collider luminosity.** In addition to energy collider has another figure of merit – its luminosity. Goal of any collider is to collide particles and to investigate products of these collisions. The productivity of the collider is defined by how effective it is in producing such collision – it is called luminosity.

As you should know, each event of interest – such as creation of new particles - is described by a cross-section, $\sigma$. By definition, the cross-section is defined in the rest frame of one type of the particles and is a constant defined by the process of interest (for example creation of Higgs boson).

Let's two species of particles with densities $n_{1,2}$ (number of particles per unit volume) and velocities $\vec{\mathrm{v}}_{1,2}$ colliding with each other. Let's consider this process in the rest frame of particle of second type (target) and the other impending this target. The number of events generated in volume $dV$ and interval time $dt$ is defined by a simple formula (coming from definition of $\sigma$):

$$d\nu = \sigma \cdot \mathrm{v}_{rel} n_1 n_2 dVdt \tag{7}$$

where $\mathrm{v}_{rel}$ is relative particle's velocity in this frame. In arbitrary (for example lab-) frame or reference we would have

$$d\nu = An_1 n_2 dVdt \tag{8}$$

where we need to define $A$. Since number of created particles (events) $d\nu$ as well as 4-volumen $dVd$ are relativistic invariants, so should be $An_1 n_2$.

We know that number of particles in a volume is also invariant, e.g.

$$ndV = n_o dV_o \rightarrow \frac{n}{n_o} = \gamma = \frac{\mathrm{E}}{mc^2}; \tag{9}$$

where index $_o$ indicated the rest frame of the particles. Hence,

$$\frac{A\mathrm{E}_1\mathrm{E}_2}{c^2 p_1^{\ i} p_{2i}} = \frac{A\mathrm{E}_1\mathrm{E}_2}{\mathrm{E}_1\mathrm{E}_2 - c^2 \vec{p}_1\vec{p}_2} = \mathrm{inv} \tag{10}$$

In the rest frame of the "target" (second type of particles)

$$\mathrm{E}_2 = m_2c^2; \vec{p}_2 = 0; \ \frac{A\mathrm{E}_1\mathrm{E}_2}{c^2 p_1^{\ i} p_{2i}} = \frac{A\mathrm{E}_1\mathrm{E}_2}{\mathrm{E}_1\mathrm{E}_2} = A = \sigma \cdot \mathrm{v}_{rel} \tag{11}$$

The same is true in the rest frame of "beam 1". Thus, in arbitrary system

$$A = \sigma \cdot \mathrm{v}_{rel} \frac{c^2 p_1^{\ i} p_{2i}}{\mathrm{E}_1\mathrm{E}_2} \tag{12}$$

Now let's look at it in the "target 2" rest frame:

$$\mathrm{E}_1 = \frac{m_1c^2}{\sqrt{1-\beta_{rel}^{\ 2}}}; \beta_{rel} = \frac{\mathrm{v}_{rel}}{c};$$

$$\frac{p_1^{\ i} p_{2i}}{c^2} = \frac{\mathrm{E}_1\mathrm{E}_2}{c^4} = \frac{m_1m_2}{\sqrt{1-\beta_{rel}^{\ 2}}} \rightarrow \beta_{rel} = \sqrt{1-\left(\frac{m_1m_2c^2}{p_1^{\ i} p_{2i}}\right)^2}; \tag{13}$$

$$\mathrm{v}_{rel} = c\sqrt{1-\left(\frac{m_1m_2c^2}{p_1^{\ i} p_{2i}}\right)^2}$$

We can express in arbitrary frame $p_1^{\,i} p_{2i}$ using 3D beam velocities

$$\frac{p_1^{\,i} p_{2i}}{m_1 m_2 c^2} = \gamma_1 \gamma_2 \left(1 - \vec{\beta}_1 \vec{\beta}_2\right) \equiv \frac{1 - \vec{\beta}_1 \vec{\beta}_2}{\sqrt{\left(1 - \vec{\beta}_1^{\,2}\right)\left(1 - \vec{\beta}_2^{\,2}\right)}}; \tag{14}$$

and after vector manipulations get final expression for $\mathrm{v}_{rel}$:

$$\frac{\mathrm{v}_{rel}}{c} = \sqrt{1 - \frac{\left(1 - \vec{\beta}_1^{\,2}\right)\left(1 - \vec{\beta}_2^{\,2}\right)}{\left(1 - \vec{\beta}_1 \vec{\beta}_2\right)^2}} = \frac{\sqrt{\left(1 - \vec{\beta}_1 \vec{\beta}_2\right)^2 - \left(1 - \vec{\beta}_1^{\,2}\right)\left(1 - \vec{\beta}_2^{\,2}\right)}}{1 - \vec{\beta}_1 \vec{\beta}_2}$$

$$\left(1 - \left(\vec{\beta}_1 \vec{\beta}_2\right)\right)^2 - \left(1 - \vec{\beta}_1^{\,2}\right)\left(1 - \vec{\beta}_2^{\,2}\right) = \left(\vec{\beta}_1 - \vec{\beta}_2\right)^2 + \left(\vec{\beta}_1 \vec{\beta}_2\right)^2 - \left(\vec{\beta}_1^{\,2}\right)\left(\vec{\beta}_2^{\,2}\right);$$

$$\left(\vec{\beta}_1^{\,2}\right)\left(\vec{\beta}_2^{\,2}\right) - \left(\vec{\beta}_1 \vec{\beta}_2\right)^2 = \left[\vec{\beta}_1 \times \vec{\beta}_2\right]^2; \tag{15}$$

$$\mathrm{v}_{rel} = c \frac{\sqrt{\left(\vec{\beta}_1 - \vec{\beta}_2\right)^2 - \left[\vec{\beta}_1 \times \vec{\beta}_2\right]^2}}{1 - \vec{\beta}_1 \vec{\beta}_2}.$$

and final expression for event rate:

$$d\nu = \sigma \frac{c^2 \sqrt{\left(p_1^{\,i} p_{2i}\right)^2 - \left(m_1 m_2 c^2\right)^2}}{\mathrm{E}_1 \mathrm{E}_2} n_1 n_2 dV dt;$$

$$d\nu = \sigma c \sqrt{\left(\vec{\beta}_1 - \vec{\beta}_2\right)^2 - \left[\vec{\beta}_1 \times \vec{\beta}_2\right]^2}\, n_1 n_2 dV dt. \tag{16}$$

For head-on collision (e.g. beam velocities lay are in the same or opposite direction) we have $\left[\vec{\beta}_1 \times \vec{\beta}_2\right]$=0 and expression is simplified significantly

$$d\nu = \sigma\left|\vec{\mathrm{v}}_1 - \vec{\mathrm{v}}_2\right| n_1 n_2 dVdt. \tag{17}$$

to a simple subtraction of the beam velocities. It means that for head on collisions $\vec{\mathrm{v}}_1 = \hat{z}\left|\vec{\mathrm{v}}_1\right|; \vec{\mathrm{v}}_2 = -\hat{z}\left|\vec{\mathrm{v}}_2\right|$ we have a sum of velocities:

$$d\nu = \sigma\left(\left|\vec{\mathrm{v}}_1\right| + \left|\vec{\mathrm{v}}_2\right|\right) n_1 n_2 dVdt \tag{17}$$

Since this condition is also maximizing the c.m. energy, it is the most favorite way to collide the beams, which we will use. Furthermore, in collider we nearly always collide ultra-relativistic particles and the sum of velocities is simply *2c*. The integral number of the events accumulated during the time period T is given by simple integral and is given by the product of the event cross-section and so-called integral luminosity:

$$\begin{gathered} \mathrm{N}(T) = \sigma\int_0^T dt \int_V \left(\left|\vec{\mathrm{v}}_1\right| + \left|\vec{\mathrm{v}}_2\right|\right) n_1 n_2\, dV = \sigma\int_0^T L(t)\,dt; \\ L(t) \underset{def}{=} \int_V \left(\left|\vec{\mathrm{v}}_1\right| + \left|\vec{\mathrm{v}}_2\right|\right) n_1(\vec{r},t)\, n_2(\vec{r},t)\, dV; \end{gathered} \tag{18}$$

where the integral is taken by the volume containing colliding beams.

Let's consider bunch trains of particles coming into the detector with frequency of $f_c$:

$$n_1(\vec{r},t) = \sum_m N_1 f_1\left(\vec{r}_\perp, s, t - \frac{m}{f_c} - \frac{s}{\mathrm{v}_1}\right); n_1(\vec{r},t) = \sum_m N_2 f_2\left(\vec{r}_\perp, s, s + \mathrm{v}_2 t, t - \frac{m}{f_c}\right) \tag{19}$$

where $N_1$ and $N_2$ are particle's number per bunch and $f_{1,2}$ are bunches distribution functions normalized to unity

$$\int f_{1,2}\, dV = 1$$

*Note that for simplicity we neglected angular spread of particles and spread in their energies – it can be taken into account for specifics of the generated products, but these details are important for detectors, not for the collider luminosity per se.*

Hence, the instant luminosity is given by convolution of the entire trains

$$L(t) = N_1 N_2 \left(|\vec{\mathrm{v}}_1| + |\vec{\mathrm{v}}_2|\right) \int_V \sum_{m,k} f_1(\vec{r}, t - mT_c) f_2(\vec{r}, t - kT_c) dV;\ T_c = 1/f_c. \tag{20}$$

while in practice only bunches with the same time stamp $k=m$ do collide in the detector, in other words the limited volume of the detector is selecting colliding bunches

$$\begin{gathered} \int_V \sum_{m,k} f_1(\vec{r}, t - mT_c) f_2(\vec{r}, t - kT_c) dV = \sum_m \int_V f_1(\vec{r}, t - mT_c) f_2(\vec{r}, t - mT_c) dV; \\ L(t) = \sum_m L_s(t - mT_c);\ L_s(t) = N_1 N_2 \left(|\vec{\mathrm{v}}_1| + |\vec{\mathrm{v}}_2|\right) \int_V f_1(\vec{r}, t) f_2(\vec{r}, t) dV; \end{gathered} \tag{21}$$

where we defined luminosity of single bunch collision, $L_s$ and assume that number of colliding particles is the same in all bunches (otherwise we need to add $N_{1m} N_{2m}$).

We are usually also not interested in details of the time structure of the single bunch collision, and we simply can define average luminosity as

$$\int_V \sum_{m,k} f_1(\vec{r},t-mT_c) f_2(\vec{r},t-kT_c)\,dV = \sum_m \int_V f_1(\vec{r},t-mT_c) f_2(\vec{r},t-mT_c)\,dV;$$

$$\bar{L}(t) = \frac{1}{T_c}\int_0^{T_c} L(t+\tau)\,d\tau \equiv \frac{1}{T_c}\int_0^{T_c} L_s(t)\,dt = \frac{f_c \cdot N_1 N_2}{\mathbf{A}} \tag{22}$$

$$\mathbf{A}^{-1} = \int_0^{T_c}\int_V \left(|\vec{\mathrm{v}}_1| + |\vec{\mathrm{v}}_2|\right) f_1(\vec{r},t) f_2(\vec{r},t)\,dV\,dt$$

where we defined the effective transverse are of the beam, $\mathbf{A}$ . By definition $f_{1,2}$ have dimensionality of inverse volume, $L^{-3}$, and integral of the product ($L^{-6}$) over the volume and *vdt* gives $L^{+4}$, which makes the dimensionality of last integral in (22) be $L^{-2}$. Hence, units for measuring the collider luminosity are in $cm^{-2}\ sec^{-1}$. Product with cross-section ($L^2$) naturally gives rate of event per second.

Let's calculate luminosity for Gaussian distribution of the beams assuming that there is no focusing elements in the collision area (typical for detectors) and $\beta$-functions have waist is the center of the detector, s=0:

$$\beta_{kx,y}(s) = \beta^*_{kx,y} + \frac{s^2}{\beta^*_{kx,y}}; k = 1,2;\ \ \sigma_{kx,y}(s) = \sqrt{\varepsilon_{kx,y}\beta_{kx,y}(s)};$$

$$f_k = \frac{1}{(2\pi)^{3/2}\sigma_{kx}\sigma_{ks}\sigma_{ks}} \exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_{kx}{}^2} + \frac{y^2}{\sigma_{ky}{}^2} + \frac{(s \pm \mathrm{v}_k t)^2}{\sigma_{ks}{}^2}\right)\right). \tag{23}$$

We will first execute integration in x and y:

$$I_x = \int_{-\infty}^{\infty} dx \exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_{1x}^{\ 2}} + \frac{x^2}{\sigma_{2x}^{\ 2}}\right)\right) = (2\pi)^{1/2}\sigma_x;\ \sigma_x^{\ -2} = \sigma_{1x}^{\ -2} + \sigma_{2x}^{\ -2};$$

$$\sigma_x = \frac{\sigma_{1x}\sigma_{2x}}{\sqrt{\sigma_{1x}^{\ 2}+\sigma_{2x}^{\ 2}}} \rightarrow \frac{1}{2\pi\sigma_{1x}\sigma_{2x}} I_x = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_{1x}^{\ 2}+\sigma_{2x}^{\ 2}}}$$

$$-//- \quad \frac{1}{2\pi\sigma_{1y}\sigma_{2y}} I_y = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_{1y}^{\ 2}+\sigma_{2y}^{\ 2}}} \tag{24}$$

$$\int dx\,dy \rightarrow \frac{1}{2\pi\sigma_{1x}\sigma_{2x}}\frac{1}{2\pi\sigma_{1y}\sigma_{2y}} I_x I_y = \frac{1}{2\pi\sqrt{\varepsilon_{1x}\beta_{1x}+\varepsilon_{2x}\beta_{2x}}\sqrt{\varepsilon_{1y}\beta_{1y}+\varepsilon_{2y}\beta_{2y}}}.$$

Note that $\beta$-functions are functions of the $s$ but not $t$ and we can make one more integral, assuming that bunch tails do not extend beyond the time between collisions:

$$\int \exp\left(-\frac{1}{2}\left(\frac{(s-\mathrm{v}_1 t)^2}{\sigma_{1s}^{\ 2}} + \frac{(s+\mathrm{v}_2 t)^2}{\sigma_{2s}^{\ 2}}\right)\right) dt; \sigma_s^{\ 2} = \frac{\sigma_{1s}^{\ 2}\sigma_{2s}^{\ 2}}{\sigma_{1s}^{\ 2}+\sigma_{2s}^{\ 2}}; \sigma_t^{\ 2} = \frac{\sigma_{1s}^{\ 2}\sigma_{2s}^{\ 2}}{\mathrm{v}_2^{\ 2}\sigma_{1s}^{\ 2}+\mathrm{v}_1^{\ 2}\sigma_{2s}^{\ 2}};$$

$$\frac{(s-\mathrm{v}_1 t)^2}{\sigma_{1s}^{\ 2}} + \frac{(s+\mathrm{v}_2 t)^2}{\sigma_{2s}^{\ 2}} = \frac{s^2}{\sigma_s^{\ 2}} + \frac{t^2}{\sigma_t^{\ 2}} - 2st\left(\frac{\mathrm{v}_1}{\sigma_{1s}^{\ 2}} - \frac{\mathrm{v}_2}{\sigma_{2s}^{\ 2}}\right);$$

$$\frac{t^2}{\sigma_{2t}^{\ 2}} - 2st\left(\frac{\mathrm{v}_1}{\sigma_{1s}^{\ 2}} - \frac{\mathrm{v}_2}{\sigma_{2s}^{\ 2}}\right) = \frac{(t-as)^2}{\sigma_t^{\ 2}} - s^2\frac{a^2}{\sigma_t^{\ 2}};\ a = \sigma_t^{\ 2}\left(\frac{\mathrm{v}_1}{\sigma_{1s}^{\ 2}} - \frac{\mathrm{v}_2}{\sigma_{2s}^{\ 2}}\right) = \frac{\mathrm{v}_1\sigma_{2s}^{\ 2} - \sigma_{1s}^{\ 2}\mathrm{v}_2}{\mathrm{v}_1^{\ 2}\sigma_{2s}^{\ 2}+\mathrm{v}_2^{\ 2}\sigma_{1s}^{\ 2}}; \tag{25}$$

$$g_s = \frac{1}{2\pi\sigma_{1s}\sigma_{2s}}\int \exp\left(-\frac{1}{2}\left(\frac{(s-\mathrm{v}_1 t)^2}{\sigma_{1s}^{\ 2}} + \frac{(s+\mathrm{v}_2 t)^2}{\sigma_{2s}^{\ 2}}\right)\right) dt = \frac{1}{\sqrt{2\pi}\tilde{\sigma}_s}\exp\left(-\frac{s^2}{2\tilde{\sigma}_s^{\ 2}}\right);$$

$$\tilde{\sigma}_s^{\ 2} = \frac{\mathrm{v}_2^{\ 2}\sigma_{1s}^{\ 2}+\mathrm{v}_1^{\ 2}\sigma_{2s}^{\ 2}}{(\mathrm{v}_1+\mathrm{v}_2)^2};\ \int g_s\,ds = 1$$

Hence, the remaining integral is

$$\mathbf{A}^{-1} = \frac{1}{2\pi\sqrt{\varepsilon_{1x}\beta^*_{1x} + \varepsilon_{2x}\beta^*_{2x}}\sqrt{\varepsilon_{1y}\beta^*_{1y} + \varepsilon_{2y}\beta*_{2y}}} \int g(s)\,ds = \frac{1}{4\pi\sigma^*_x\sigma^*_y} h\left(\beta^*_{1,2x,y}, \sigma_s\right)$$

$$\sigma^{*\,2}_{x,y} = \frac{\varepsilon_{1x,y}\beta^*_{1x,y} + \varepsilon_{2x,y}\beta^*_{2x,y}}{2};\ h\left(\beta^*_{1,2x,y}, \sigma_s\right) = \int g(s)\,ds;\ \ g(s) = g_x g_y g_s;$$

$$g_{x,y} = \sqrt{\frac{\varepsilon_{1x,y}\beta^*_{1x,y} + \varepsilon_{2x,y}\beta^*_{2x,y}}{\varepsilon_{1x,y}\beta_{1x,y}(s) + \varepsilon_{2x,y}\beta_{2x,y}(s)}} = \left(1 + \frac{s^2}{\tilde{\beta}^{*2}_{x,y}}\right)^{-1/2} = g_{x,y}\left(\frac{s}{\tilde{\beta}^*_{x,y}}\right) \tag{26}$$

$$\tilde{\beta}^{*2}_{x,y} = \beta^*_{1x,y}\beta^*_{2x,y} \cdot \frac{\varepsilon_{1x,y}\beta^*_{1x,y} + \varepsilon_{2x,y}\beta^*_{2x,y}}{\varepsilon_{2x,y}\beta^*_{1x,y} + \varepsilon_{1x,y}\beta^*_{2x,y}};$$

$$\overline{L}(t) = \frac{f_c \cdot N_1 N_2}{4\pi\sigma^*_x\sigma^*_y} h\left(\frac{\beta^*_{x,y}}{\tilde{\sigma}_s}\right).$$

where we introduce so called hour-glass effect function which takes into account length of the bunches. Carrying out some parameters, which is overestimation of accuracy is indeed unnecessary and we can use

$$\left|\vec{\mathrm{v}}_1\right| = \left|\vec{\mathrm{v}}_2\right| = c;\ \sigma_s = c\sigma_t = \sqrt{2}\tilde{\sigma}_s;$$

$$g_s = \frac{1}{\sqrt{2\pi\left(\sigma_{1s}^{\ 2} + \sigma_{2s}^{\ 2}\right)}} \exp\left(-\frac{s^2}{\sigma_{2s}^{\ 2} + \sigma_{1s}^{\ 2}}\right); \tag{27}$$

The first term in expression for luminosity is its maximum value, which could be achieved for a very short bunches:

$$\tilde{\beta}^*_{x,y} >> \tilde{\sigma}_s; g_{x,y} \cong 1; \int g(s)\,ds = 1$$

$$\sigma^{*\ 2}_{x,y} = \frac{\varepsilon_{1x,y}\beta^*_{1x,y} + \varepsilon_{2x,y}\beta^*_{2x,y}}{2}; \tag{28}$$

$$\bar{L}(t) \cong \frac{f_c \cdot N_1 N_2}{4\pi\sigma^*_x\sigma^*_y} = \frac{f_c \cdot N_1 N_2}{2\pi\sqrt{\varepsilon_{1x}\beta^*_{1x} + \varepsilon_{2x}\beta^*_{2x}}\sqrt{\varepsilon_{1y}\beta^*_{1y} + \varepsilon_{2y}\beta*_{2y}}}$$

with even more simple formulae for beam with equal emittances and $\beta$-functions:

$$\beta^*_{1,2x,y} = \beta^*_{x,y}; \varepsilon_{1,2x,y} = \varepsilon_{x,y} \Rightarrow \bar{L}(t) = \frac{f_c \cdot N_1 N_2}{4\pi\sqrt{\varepsilon_x\beta^*_x}\sqrt{\varepsilon_y\beta^*_y}} h \tag{29}$$

or round beam (as in typical hadron colliders):

$$\beta^*_{x,y} = \beta^*; \varepsilon_{x,y} = \varepsilon \Rightarrow \bar{L}(t) = \frac{f_c \cdot N_1 N_2}{4\pi\varepsilon\beta^*} h\left(\beta^*, \sigma_s\right). \tag{30}$$

Thus, one can increase luminosity buy increasing frequency of collision (e.g. increasing average beam currents $I_1, I_2, L \sim \sqrt{I_1 I_2}$), increasing number of particles per bunch (more effective way since $L \sim I_1 I_2$), deducing $\beta^*$ (can be limited by beam optics or by bunch lengths, $\sigma_s$) or reducing emittance(s).

**Beam-beam effects and limits.** Two colliding beams sample strongly non-linear transverse EM fields induced by the opposite beam. It is very important to observe that particles also sample EM fields generated by their own bunch, but its effect is relativistically suppressed. This can be clearly demonstrated in a following way: moving particle generates transverse electric and magnetic field (check your E&M)

$$\vec{B}_\perp = \frac{\text{v}}{c}[\hat{z}\times\vec{E}_\perp];\vec{\text{v}}=\hat{z}\text{v} \tag{31}$$

and a particle moving in the same direction and with the same velocity experiencing Lorentz force of:

$$\begin{aligned}
&\frac{d\vec{p}_\perp}{dt} = e\left(\vec{E}_\perp + \left[\frac{\vec{\text{v}}}{c}\times\vec{B}_\perp\right]\right) = \vec{E}_\perp + \left(\frac{\text{v}}{c}\right)^2\left[\hat{z}\times[\hat{z}\times\vec{E}_\perp]\right] \\
&\left[\hat{z}\times[\hat{z}\times\vec{E}_\perp]\right] = \hat{z}\underbrace{\left(\hat{z}\vec{E}_\perp\right)}_{0} - \vec{E}_\perp\left(\hat{z}\hat{z}\right) = -\vec{E}_\perp; \\
&\frac{d\vec{p}_\perp}{dt} = e\vec{E}_\perp\left(1-\left(\frac{\text{v}}{c}\right)^2\right) = \frac{e\vec{E}_\perp}{\gamma^2}
\end{aligned} \tag{32}$$

which is relativistically suppressed by huge factor $\gamma^2$. Note that for colliders $\gamma \propto 10^2 - 10^4$ are typical. In contrast, particle on the colliding course $\vec{\text{v}}=-\hat{z}\text{v}$ experiences Lorentz force of

$$\frac{d\vec{p}_\perp}{dt} = e\left(\vec{E}_\perp - \left[\frac{\vec{\text{v}}}{c}\times\vec{B}_\perp\right]\right) = \vec{E}_\perp\left(1+\left(\frac{\text{v}}{c}\right)^2\right) \tag{33}$$

which for ultra relativistic particles is simply doubles that of electric field. Since these EM fields are generated by the beams themselves, they are strongly non-linear with typical scale of the filed variations defined by the transverse beam size.

Electric field of a Gaussian bunch can be found in its co-moving frame and then transferred into the lab frame using Lorentz transformation for the EM fields. In the co-moving the bunch length is increasing by the factors of $\gamma$ . While even in the lab-frame bunch length is much larger than its transverse size, it is definitely true in the beam-frame for all operating colliders. Thus, we have to solve Poisson equation for

$$\Delta\varphi = -4\pi\rho; \rho = \frac{\rho_o}{2\pi\sigma_x\sigma_y}\exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_x}+\frac{y^2}{\sigma_y}\right)\right) \tag{34}$$

which can be solve using Fourier transform:

$$\iint ..\exp\left(\vec{k}\vec{r}\right)dxdy/(2\pi)^2 \to \varphi\left(\vec{k}\right) = 4\pi\rho\left(\vec{k}\right)/\vec{k}^2$$

$$\frac{1}{\vec{k}^2} = \int_0^\infty e^{-\vec{k}^2 t}dt \equiv \frac{1}{4}\int_0^\infty e^{-\frac{\vec{k}^2 t}{4}}dt$$

and scaling it by 1/4th we get:

$$\varphi\left(\vec{k}\right) = 4\pi\int_0^\infty \rho\left(\vec{k}\right)e^{-\vec{k}^2 t}dt;\ \varphi\left(\vec{r}\right) = \pi\int_0^\infty dt\iint e^{-i\vec{k}\vec{r}}\rho\left(\vec{k}\right)e^{-\frac{\vec{k}^2 t}{4}}dk_x\,dk_y.$$

Then, for a long Gaussian bunch with linear density of $\rho_o\left(z\right) = eZN\cdot e^{-\frac{z^2}{2\bar{\sigma}_z^2}}/\left(\sqrt{2\pi}\bar{\sigma}_z\right)$.

$$\rho = \rho_o\left(z\right)\frac{1}{2\pi\sigma_x\sigma_y}e^{-\frac{x^2}{2\sigma_x^2}-\frac{y^2}{2\sigma_y^2}};\quad \rho\left(k\right) = \frac{1}{\left(2\pi\right)^2}\cdot e^{-\frac{k_x^2\sigma_x^2}{2}-\frac{k_y^2\sigma_y^2}{2}};$$

after trivial integration,

$$\varphi\left(\vec{r}\right) = \pi\rho_o\left(z\right)\frac{1}{\left(2\pi\right)^2}\cdot\int_0^\infty dt\iint e^{-i\vec{k}\vec{r}}e^{-\frac{k_x^2\sigma_x^2}{2}-\frac{k_y^2\sigma_y^2}{2}}e^{-\frac{\vec{k}^2 t}{4}}dk_x\,dk_y;\quad \int_{-\infty}^{\infty}e^{-ik_x x}e^{-\frac{k_x^2\left(2\sigma_x^2+t\right)}{4}}dx = \sqrt{\frac{4\pi}{2\sigma_x^2+t}}e^{-\frac{x^2}{2\sigma_x^2+t}};$$

we get the desirable result identical to (A7):

$$\varphi\left(\vec{r}\right) = \frac{eZN}{\sqrt{2\pi}\bar{\sigma}_z}e^{-\frac{z^2}{2\sigma_z^2}}\cdot\int_0^\infty \frac{e^{-\frac{x^2}{2\sigma_x^2+t}-\frac{y^2}{2\sigma_y^2+t}}}{\sqrt{\left(2\sigma_x^2+t\right)\left(2\sigma_y^2+t\right)}}dt. \tag{35}$$

where t is integration parameter, not the time!

What is true about (35) is that the field is indeed very nonlinear and effect on colliding particles is reducing when they go outside of the beam core. Expansion of the potential near the beam axis gives a simple expression:

$$\bar{\varphi}(x,y,\bar{z}) = -\frac{eN}{\sqrt{2\pi}\bar{\sigma}_z} e^{-\frac{\bar{z}^2}{2\bar{\sigma}_z^2}} \cdot \frac{2}{\sigma_x+\sigma_y}\left(\frac{x^2}{\sigma_x}+\frac{y^2}{\sigma_y}\right) + O^4 . \tag{36}$$

with electric field simply calculated by

$$\vec{E}_{bf} = -\vec{\nabla}\bar{\varphi}(x,y,\bar{z}) = \frac{eN}{\sqrt{2\pi}\bar{\sigma}_z} e^{-\frac{\bar{z}^2}{2\bar{\sigma}_z^2}} \cdot \left(\hat{x}\frac{x}{\sigma_x(\sigma_x+\sigma_y)} + \hat{y}\frac{y}{\sigma_y(\sigma_x+\sigma_y)}\right) \tag{37}$$

e.g. the beams of the same charge sign will experience defocusing (repulsion, for example *pp*) and beams of the opposite charge signs will experience focusing (for example in electron-position collides). Lorentz transferring the field into the lab frame and applying the formulae for the tune shifts we already derived we have (q is the charge of the particle in the opposite beam):

$$\vec{E} = \gamma \vec{E}_{bf}; \vec{B} = \gamma\beta\left[\hat{z}\times\vec{E}_{bf}\right];$$

$$\Delta Q_{2x,y} = -\frac{qeN_1}{2\pi\gamma m_2 c^2}\int \frac{\beta_{2x,y}}{\sigma_{1x,y}(\sigma_{1x}+\sigma_{1y})} e^{-\frac{s^2}{2\sigma_{1s}^2}} \frac{ds}{\sqrt{2\pi}\sigma_{1s}}; \tag{38}$$

$$\Delta Q_{1x,y} = -\frac{qeN_2}{2\pi\gamma m_1 c^2}\int \frac{\beta_{1x,y}}{\sigma_{2x,y}(\sigma_{2x}+\sigma_{2y})} e^{-\frac{s^2}{2\sigma_{2s}^2}} \frac{ds}{\sqrt{2\pi}\sigma_{2s}};$$

For a short bunches, $\sigma_s << \beta_{x,y}$ we can move the expression from the integral and have an approximation for the beam-beam tune shifts:

$$\Delta Q_{2x,y} \cong -\frac{qeN}{2\pi\gamma m_2 c^2}\frac{\beta^*_{2\ x,y}}{\sigma^*_{1x,y}\left(\sigma^*_{1x}+\sigma^*_{1y}\right)} \ \& \ 1 \Leftrightarrow 2 \qquad (39)$$

Considering again round beams with the same $\beta$-functions and emittance, we would have bema-beam tune shifts of:

$$\Delta Q_{1,2} = -\frac{qeN_{2,1}}{4\pi\gamma m_{1,2}c^2}\int\frac{\beta(s)}{\varepsilon\beta(s)}e^{-\frac{s^2}{2\sigma_s^2}}\frac{ds}{\sqrt{2\pi}\sigma_s} = -\frac{qeN_{2,1}}{4\pi\gamma m_{1,2}c^2\varepsilon}$$
$$\Delta Q_{1,2} == \pm\frac{N_{2,1}r_{1,2}}{4\pi\gamma\varepsilon};\ r_{1,2} = \frac{e^2}{m_{1,2}c^2}. \qquad (40)$$

Note that the beam-beam tune shift is inverse proportional to emittance.

Since the tune shift is reducing with the amplitude (particles far away from the beam see field ~ 1/r), the bema-beam tune tuned shift is equal to the tune spread in the beam. Hence, some of the particles can get to strong resonance and get lost (or move to large amplitudes). Otherwise, which frequently happens in 6D phase space, a stochastic layer occurs and particle diffuse to large amplitudes, the beam emittances and sized increase and luminosity is reduced.

In any case, accelerator physicist tried multiple tricks (including colliding 4 bunches with two electron-position beams) and found that there is practical limitation for ring-ring colliders:

1. For hadron ring where damping is absent, maximum achievable tune shifts are $|\Delta Q_h| \le 0.02$;
2. For lepton colliders with strong damping (~ 1,000 turns) $|\Delta Q_l| \le 0.1$.

It is easy to make it worth, but so far nobody manage to exceed these limits. Thus, both reducing beam emittances and increasing the number particles per bunch can be used for increasing luminosity when tune shifts do not exceed the limits imposed by the beam dynamics.

**Nonlinear resonances**. Let's examine why this is happening? We start from considering a single resonant term resulting from nonlinearity in a one term map:

$$\mathcal{M} = exp(-:\boldsymbol{\mu}\cdot\mathbf{I} + D(\mathbf{I}):)exp\left(:f_3:\right)exp\left(:f_4:\right)\ldots$$

with higher order terms containing resonance driving terms

$$h_{nmk} I_1^{\frac{|m|}{2}} I_2^{\frac{|n|}{2}} I_3^{\frac{|k|}{2}} \cos\left(m\varphi_1 + n\varphi_2 + k\varphi_3\right) \tag{41}$$

at resonance

$$m\mu_1 + n\mu_2 + k\mu_3 = 2\pi N$$

The meaning of the resonant conditions that even a very weak resonant term plays important role at exact resonance because it applies turn after turn with the same value, i.e. its effect accumulates while non-resonant terms oscillate and average to zero. This why resonant term can dominate when conditions are sufficiently close to the resonance.

While it will be interesting to discuss coupling resonance, it is much easier to discuss and, even more importantly, to plot one-dimensional resonances. Let's do exactly this and consider a 1D map with a single resonance

$$m\mu = N + \delta\mu;\ |\delta\mu| << 1$$

where $\delta\mu$ is detuning from the resonance. The corresponding Hamiltonian can be written as

$$H(\varphi, I) = \delta\mu\cdot I + D(I) + h_m(I)\cdot\cos m\varphi;\ \langle osc.terms\rangle = 0 \tag{42}$$

It means that we can find stationary points beyond the origin *I=0*:

$$I' = -\frac{\partial H}{\partial \varphi} = mh_m(I)\cdot \sin m\varphi = 0 \Rightarrow m\varphi_r = M\pi;$$

$$\varphi' = \frac{\partial H}{\partial I} = \delta\mu + \frac{\partial D}{\partial I} + \frac{\partial h_m}{\partial I}\cdot \cos m\varphi_r \equiv \delta\mu + \frac{\partial D}{\partial I} \pm \frac{\partial h_m}{\partial I} = 0, \tag{43}$$

with half of the stationary point being stable and half unstable (separatrix crossings). Since we assume that resonance is weak, we can expand Hamiltonian near stationary points:

$$m\varphi = m\,\Delta\varphi;\quad I = I_+ + \Delta I;\ \ \delta\mu + \partial_I D(I_+) + \partial_I h_m(I_+) = 0$$

$$m\varphi = \pi + m\,\Delta\varphi;\quad I = I_- + \Delta I;\ \ \delta\mu + \partial_I D(I_-) - \partial_I h_m(I_-) = 0$$

$$H_\pm = \pm h_m(I_\pm)\left(1 - \frac{m^2\Delta\varphi^2}{2}\right) + \left(\partial^2{}_I D(I_\pm) \pm \partial^2{}_I h_m(I_\pm)\right)\frac{\Delta I^2}{2}$$

and when second derivatives in the second term do not change sign (again, assuming $h_m$ is small), one of the phases (odd or even) is table and corresponds to oscillations inside resonance's separatrix. The other phase position is unstable and corresponds to separatrix crossing point. Without loss of generality, let's assume that it looks as in Fig. 5.

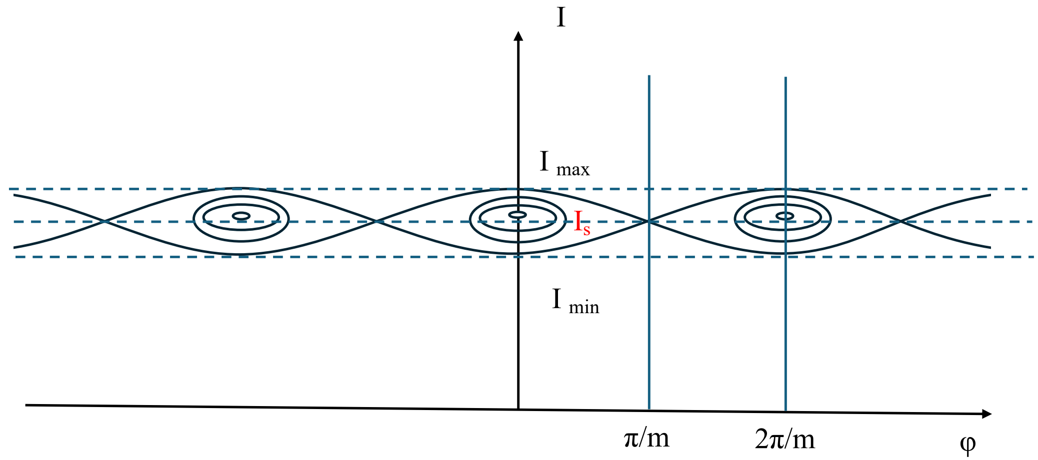



Fig. 5. Phase space for a generic isolated 1D resonance

We can estimate width of the resonance, i,e. distance between $I_{max}$ and $I_{min}$ by understanding that Hamiltonian is constant long the separatrix and

$$\left(\partial^2{}_I D \pm \partial^2{}_I h_m\right)\frac{\Delta I^2}{2} = 2\left|h_m\right| \rightarrow \Delta I = \pm\sqrt{\frac{\left|h_m\right|}{\left|\partial^2{}_I D - \partial^2{}_I h_m\right|}};\ I_{max} - I_{min} = 2\sqrt{\frac{\left|h_m\right|}{\left|\partial^2{}_I D - \partial^2{}_I h_m\right|}} \quad (44)$$

where values are taken at $I=I_s$. More accurate estimation will be from the using Hamiltonian (52)

$$H(0,I_m) = H(\pi,I_s) \Rightarrow \delta\mu \cdot I_m + D(I_m) + h_m(I_m) = \delta\mu \cdot I_s + D(I_s) - h_m(I_s);$$

$$\frac{D''(I_s) + h''_m(I_s)}{2}(I_m - I_s)^2 + \left(\delta\mu + 2h'_m(I_s)\right)(I_m - I_s) + 2h_m(I_s) = 0;$$

$$I_m - I_s = -2\frac{\left(\delta\mu + 2h'_m(I_s)\right) + \sqrt{\left(\delta\mu + 2h'_m(I_s)\right)^2 - h_m(I_s)\left(D''(I_s) + h''_m(I_s)\right)}}{D''(I_s) + h''_m(I_s)};$$

$$I_{max} - I_{min} = 4\frac{\sqrt{\left(\delta\mu + 2h'_m(I_s)\right)^2 - h_m(I_s)\left(D''(I_s) + h''_m(I_s)\right)}}{D''(I_s) + h''_m(I_s)}$$

But what is probably more interesting is to see if structure of a weak resonant brakes in in the presence of damping or oscillations. Let's modify equation (43) by adding damping term

$$I' = mh_m(I)\cdot \sin m\varphi - \zeta \cdot I \tag{45}$$

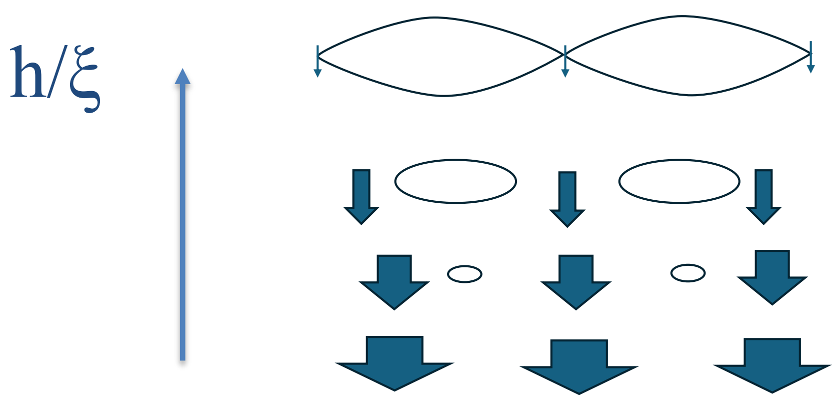
to see that stationary points can only occur only if

$$m\left|h_m(I)\right|\cdot \geq \zeta \cdot I \tag{46}$$

If this condition is not satisfied and

$$\zeta \cdot I > m\left|h_m(I)\right|$$

the resonance structure will be destroyed by damping. This is the main reason why high order resonances can be practically ignored in electron and positron storage rings where – with proper lattice - synchrotron radiation provides for strong damping of all three degrees of freedom. Naturally, this is not the case for hadron storage rings: even in LHC damping is too feeble to destring 5$^{th}$ or 6$^{th}$ order resonances.



Now we are equipped with “instruments” to discuss why nonlinear maps can result in chaotic trajectories at large amplitudes

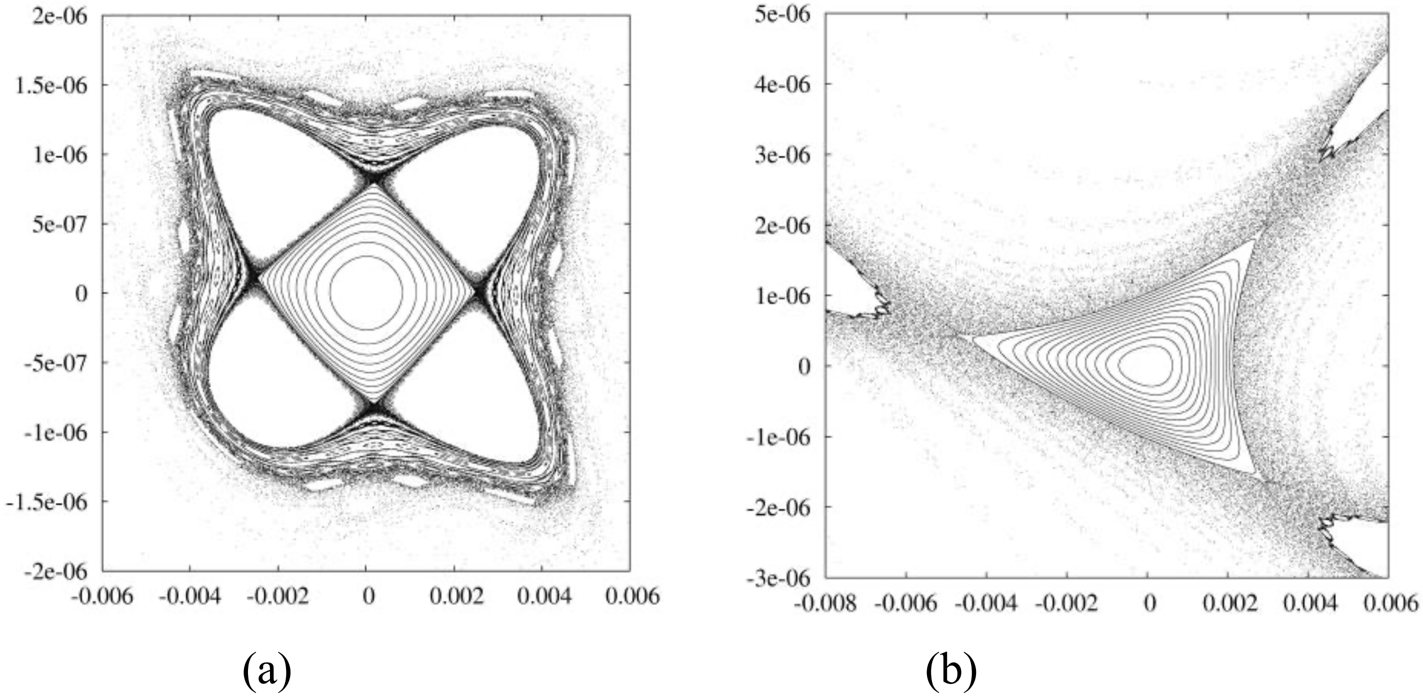



Fig.6. Two tracking results: (a) with 4th order and (b) 3rd order resonance strictures.

First, it is worth mentioning that crossing points of separatrices are intrinsically unstable – formally it takes infinite time for particle to reach them: it means that they can be easily destroyed by very weak disturbances. Indeed, this is usually points where stochastic behavior originates from. While some specific cases can be analyzed (to a degree) analytically, there is no general theory of when nonlinear system becomes chaotic.

But there is one logical – even though an approximate – criteria was developed by I. Chirikov. His hypothesis was that chaos occurs when two non-linear resonances overlap in the phase space.

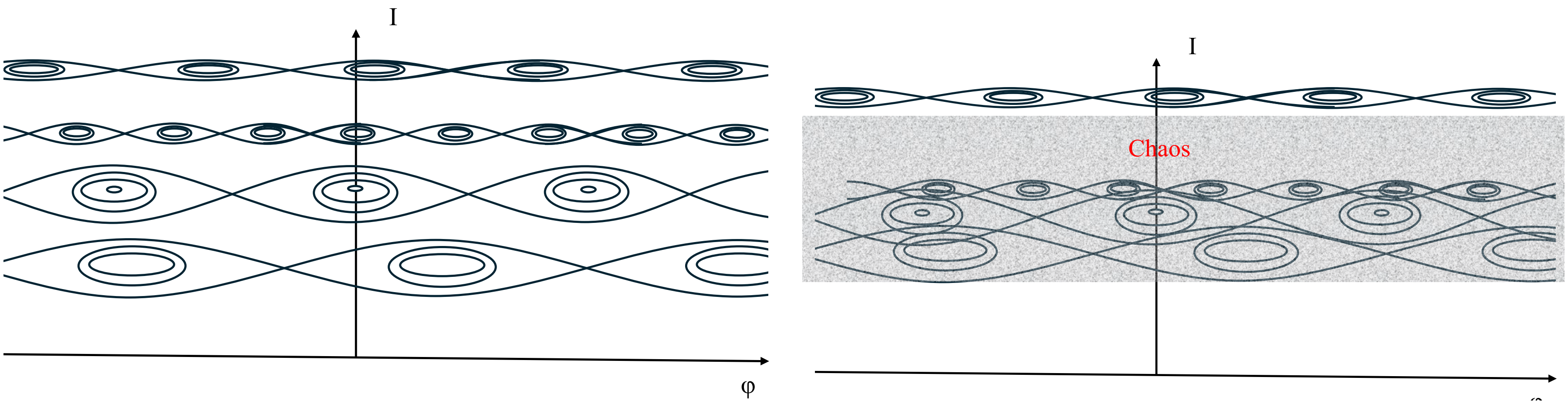
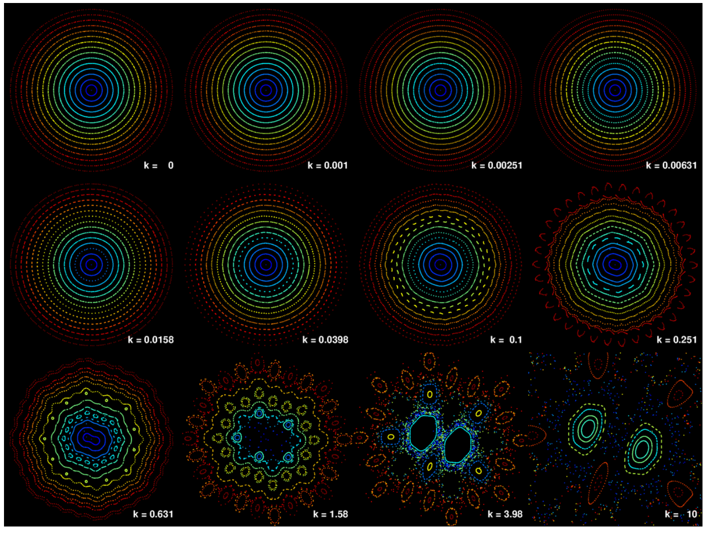



Fig.7. From regular motion to chaos as result of overlapping resonances.

Overlap of resonances can happened because of poor choice of operating point: for example, close to 3rd order resonance, which would overlap with 6th order resonance and create unstable trajectories. But even with well chosen tunes, resonances can overlap when their strength is increasing. One typical cases will in light sources operating with very low emittances: compensation of chromaticities requires very strong sextupoles, which generate resonances of all orders. Unless there a special treatment of resonant terms, the ring heating what is known “as a chromaticity brick wall” resulting in collapsing dynamic aperture: i.e. the range of amplitudes (actions) available for stable motion.

But the most classical – and probably the most important - case is the beam-beam effects in colliders. First, and foremost, beam-beam kick is completely nonlinear at the scale if the transverse beam sizes. Maximum tune shift is at the beam center, and fields are falling as $1/r^2$ at large amplitudes and practically vanished at few RMS beam sized. It means that tune spread in the beams can be approximated by the beam-beam tune shift. It also means that with tune spread ~ 0.1, the 3rd and 4th order resonances (separated by 0.083) would overlap if betatron tunes chosen in vicinity. This is why for high luminosity e+e- colliders, tunes are usually located near integer resonance: above or below depending on the value of the tune shift with amplitude. But in any case, continuous increasing beam-beam effects would result in bring chaotic boundary to the beam vicinity. At best it would result in increase in the beam emittance and reduction of the beam-beam effects and reduction of luminosity. In worst case, the beam would quickly die… Thus, we have to obey these limitations.

Hence, understanding of the nonlinear beam dynamics is critically important for designing high luminosity colliders and high brightness synchrotron radiation sources.



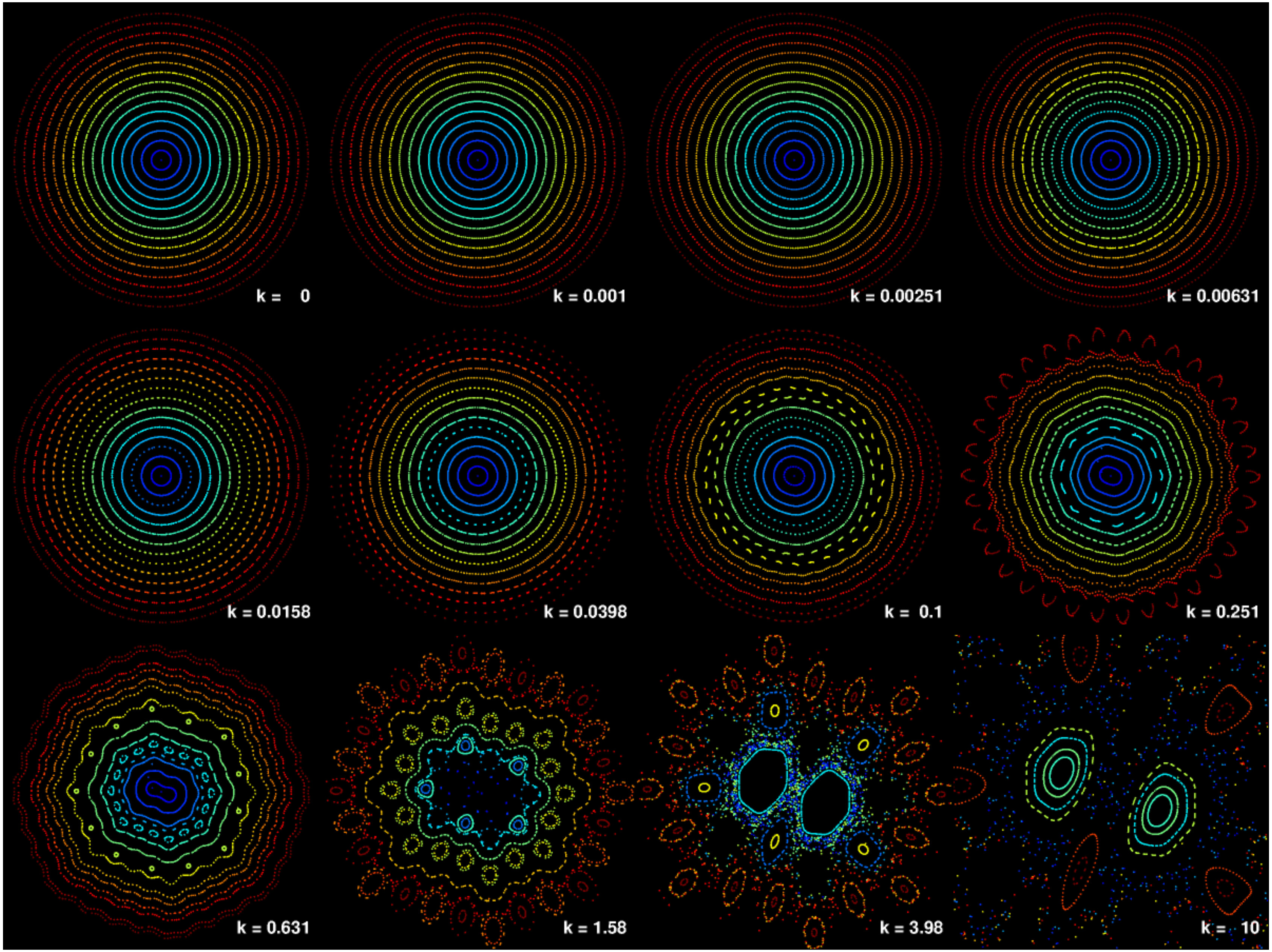



Fig. 8. Modification of the phase space structure with the increase of the beam-bema tune shift.

# What we learned today

- Colliders have two main figures of merit: center-mass energy and luminosity
- With a given beam intensities the luminosity can be increased by reducing beam emittances but this increase is limited by beam-beam effects
- Beam-beam effects are intrinsically non-linear and drive nonlinear resonances of all orders
- Weak resonance can be destroyed by radiation damping
- But strong resonances will limit maximum beam-beam tune shifts below 0.1 for e+e- colliders and below 0.02 in hardon colliders