Hyperfinecourse A: framework

February 5, 2020

Abstract

This document is meant for optional background reading when studying www.hyperfinecourse.org. It deals with one of the chapters of this course. The formal course content is defined by the website and videos. The present document does not belong to the formal course content. It covers the same topics, but usually with more mathematical background, more physical background and more examples. Feel free to use it, as long as it helps you mastering the course content in the videos. If you prefer studying from the videos only, this is perfectly fine.

The present text has been prepared by Jeffrey De Rycke (student in this course in the year 2018-2019). He started from a partial syllabus written by Stefaan Cottenier for an earlier version of this course, and cleaned, edited and elaborated upon that material. That syllabus was itself inspired by a course taught by Michel Rots at KU Leuven (roughly 1990-1995).

1 VIP-1

This section will deal with the logical build up of VIP-1. It will cover the same "ladder" model the video has, but each step will be explained more in details. We will start from how we used to describe basic atoms. Then work our way down from high energy level splitting (excited nuclei), to the energy levels of the excited electrons, to the L-S coupling (fine splitting), all the way to hyperfine splitting.

1.1 Approximations of Basic Nuclei

To describe the H (or even the He atom) we made some basic approximations to make the calculations feasible, these approximations were will be described in short.

1.1.1 Non-relativistic

The non-relativistic quantum mechanical equation for the hydrogen atom (and for any other quantum system) is the Schrödinger equation. Its relativistic analogue is the Dirac equation. To first order in v^2/c^2, the Dirac equation can be approximated by the Schrödinger equation plus three extra terms. Those terms represent the most important relativistic effects. They are:

- *The mass-velocity effect.* The dependence of the electron mass on the electron velocity causes altered orbits for high-speed electrons (= the ones closest to the nucleus).
- *The Darwin effect.* In relativistic quantum physics, the electron can be shown to execute extremely fast random movements over a short length scale, known as the *zitterbewegung* [1]. Therefore, the electron experiences the Coulomb potential by the nucleus as somewhat smeared out, which slightly changes the energy levels of the hydrogen and other atoms.
- *Spin-orbit coupling.* The orbit of an electron (and hence the energy of both the electron and the atom) does not depend only on the electromagnetic interaction of its charge with the charge of the nucleus, but also on its spin: the electron's orbital motion generates a magnetic field, that interacts with its spin magnetic moment. This interaction between spin and orbit is called spin-orbit coupling. Because spin-orbit coupling – and therefore spin – naturally shows up in the Dirac theory, spin can be understood as being a relativistic effect.

[1] This should not be confused with one of the Heisenberg uncertainty principles, which is present already in non-relativistic quantum physics.

1.1.2 Effective Electron-Electron Interactions

When having multiple electrons, solving the electron-electron interactions is exceedingly difficult. Solving the Hamiltonian directly (by expanding the many-electron wave function into a linear combination of Slater determinants) becomes unfeasible as one needs many Slater determinants when the amount of electrons grows and one wants a reasonable accurate solution.

A solution to this is identifying a subspace of the full Hilbert space in which the e-e interactions plays a decisive role. The e-e interaction is then treated explicitly within this limited subspace, while the influence of the rest of the Hilbert space is treated via a mean-field approximation. This greatly reduces the interaction terms as we let the prominent interactions happen in small spaces.

This process of describing the e-e interactions as prominent in certain subspaces (and the rest of the space via a mean-field) is what effective electron-electron interactions are comprised off.

1.1.3 Infinitely Heavy Nucleus

When the nucleus is not infinitely massive any more, electron and nucleus will move about their common centre of mass which lies close to the nucleus[2]. The orbit of the electron (therefore also its energy)will slightly change. Energies corrected for this effect are obtained by multiplying the results from the infinitely massive proton by a factor $\frac{1}{1+\frac{m_e}{M}}$ (M is the mass of the nucleus) . This increases the ground state energy of H by 0.008 eV. For Deuterium ('heavy Hydrogen', with a nucleus of 1 proton and 1 neutron) this formula tells its ground state is 0.004 eV lower than the H ground state. For really heavy nuclei where M is high, the effect of this approximation becomes progressively small.

In contrast to this classical mass effect, a second correction due to finite mass is of a purely quantumphysical origin: the zero-point motion of light objects. A quantummechanical object can never be at rest, even not at 0 K and in its absolute ground state. Always there will be a random vibration about its rest position, a phenomenon that is called zero-point motion and can be understood as a consequence of the uncertainty principle. The lighter the object, the more important the zero-point motion. Nuclei are usually sufficiently heavy to make this zero-point motion negligibly small, but for the lightest nuclei it can become significant.

[2]The proton-to-electron mass ratio is about 1836.

1.1.4 Point Nucleus

The fourth approximation we will discuss is the assumption that the nucleus has no structure. What do we mean by this? In the first place that all protons and neutrons in the nucleus coincide with the same mathematical point. If this is not the case - imagine a spherical, cigar-shaped or even more complex nucleus - the Coulomb potential due to the nucleus will not be spherically symmetric any more. This will result in other solutions and therefore energy levels of our Hamiltonian.

The same happens if the nucleus has structure in a magnetic sense: it can have a magnetic dipole moment (or even a higher order multipole moment). Because of the dipole moment the nucleus will generate a magnetic field, and this breaks the spherical symmetry and hence leads to preferred orientations of the electrons.

And now we have finally arrived at the subject of this course: the role of the structure of the nucleus, in its spatial and magnetic sense. As we will calculate later, the new influence of the nucleus we take now into account will introduce new shifts and splittings of the order of μeV. Because the energy scale which is needed to describe these new splittings is orders of magnitude smaller than the energy scale for the fine structure, the new details in the atomic spectrum are called the *hyperfine structure* of the atom.

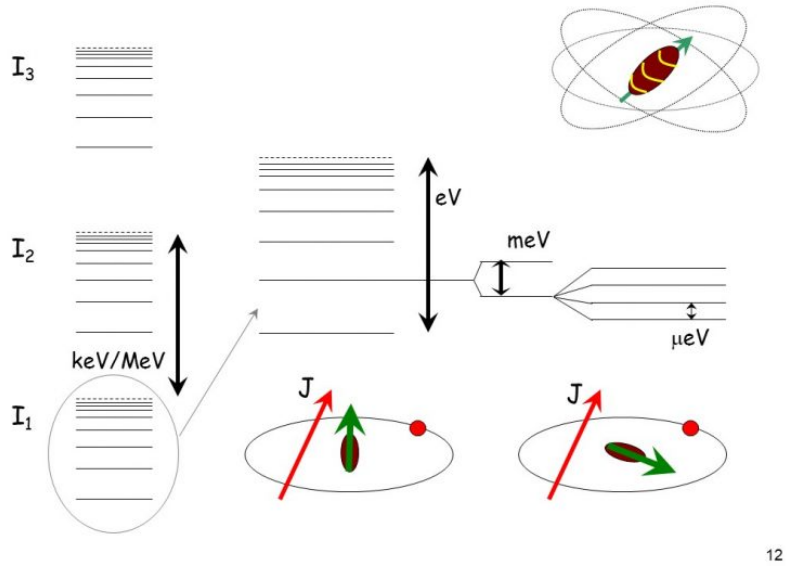

Figure 1: VIP number one. It shows the different energy levels and splittings of an atom.

1.2 The Energy Levels

While reading this section, keep an eye on Fig. 1 p.4. We will constantly be referencing the energy levels on this picture while describing every part of said image. The energy levels on the left are all possible energy levels of the "nucleus + electrons" system. The sections below describe the different attributions to said energy levels. Unfortunately it is not possible to "see" all different energy levels on said energy scale. That's why we zoom in where needed in the centre and on the right side of the image.

1.2.1 Nucleus Excitations

As described in the previous chapter, a nucleus can get excited when the internal distribution of protons and neutrons change. This results in a configuration with exactly the same particles, yet with a higher overall energy (i.e. a higher energy level). These configurations are the most left energy levels I_1, I_2,... This can be done via absorbing a photon in the keV/MeV energy range. Or when a nucleus decays to another nucleus, which can be not yet in its ground state after decaying. As seen via the energy of the photons, the difference in energies are of order keV/MeV.

1.2.2 Electron Excitations

In the Rutherford-Bohr model, electrons orbit the nucleus in well defined orbits. These well defined orbits have well defined energy levels, which are quantized, as seen on the second energy splitting (counting from the left). The difference in energy values are of order eV. It therefore also takes photons in the order of eV to excite electrons to higher orbits.

1.2.3 Spin-Orbit Coupling (fine splitting)

This is where the energy splitting starts to get more complicated. The spin-orbit coupling is a weak magnetic interaction (coupling) between the electron spin and its orbital motion. The intrinsic spin of the electron creates a spin magnetic dipole moment. While, from the restframe of the electron, the rotating nucleus creates a magnetic field. This is why it is called a relativistic effect, it arises from putting ourselves at the position of a stationary electron and moving nucleus. For light atoms, the individual spins $\mathbf{s}_i$ will interact with each other to form a total spin angular momentum $\mathbf{S}$. Likewise, the individual orbital angular momentum $\mathbf{l}_i$ form a total orbital angular momentum $\mathbf{L}$. These quantum numbers $\mathbf{S}$ and $\mathbf{L}$ interact with each other via what is called Russell-Saunders coupling (or simply LS coupling). The $\mathbf{S}$ and $\mathbf{L}$ couple together and form a total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$. This value depends on the relative orientation between $\mathbf{L}$ and $\mathbf{S}$. Just as how the energy of a system is different between different orientations of a bar magnet in a magnetic field, the energy correction is different for different values of $\mathbf{J}$ as illustrated in the figure on the next page.

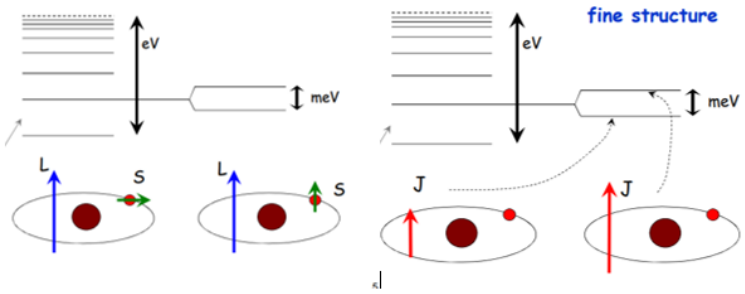

Figure 2: Coupling of different **L** and **S** into different **J**, and its effect on the energy splitting.

On the left, two different relative orientations are shown, each resulting in a different **J**. These different values result in different energy splitting of order of meV.

There is a way of characterizing different states, this is done via the term symbols. A term symbol is defined via ${}^{2S+1}L_J$. For same values of S and L, different orientations (and therefore different values of J) are possible. J ranges from L+S to |L-S| in steps of -1. When filling in L, we use the symbols S, P, D, F, G, H,... instead of 0, 1, 2,...

As stated before, different orientations will result in different energy levels. To know the relative orientation of the energy levels, we can use Hund's rules. They say the following:

- The highest multiplicity (defined as 2S+1) has the lowest energy.
- For the same multiplicity, the largest L has the lowest energy.
- For the same L and half filled or less filled shells, the lowest J has the lowest energy. For the same L and more than half filled shells, the highest J as the lowest energy.

When talking about shells, we of course mean the outer most shell. As electrons of filled shells balance each other in $\mathbf{s}_i$ and $\mathbf{l}_i$, therefore creating S = L = 0 and by extend J = 0.

Take, for example, Sodium. It has a filled s and p shell, and one electron in the 2s shell. This gives us the term ${}^2S_{1/2}$. When exciting this electron to the 2p shell, we have two possible terms. ${}^2P_{1/2}$ and ${}^2P_{3/2}$ (via S = 1/2, the same as before, and L now = 1). Where Hund's rules dictate that $\mathrm{E}({}^2P_{3/2}) > \mathrm{E}({}^2P_{1/2})$. This fine splitting of energy levels is something one can see when looking at the spectrum of a Sodium lamp. An orange doublet can be seen at wavelengths of 589.6 nm and 589.0 nm.

1.2.4 Hyperfine Splitting

Two different hyperfine splittings can be identified. The electric hyperfine splitting and the magnetic hyperfine splitting. The following two sections will be rather short, as they both get their own chapter where they will be broadly discussed.

The electric hyperfine splitting has to do with the interaction of the nuclear quadrupole moment and the electric-field gradient (of the electron cloud). The electric-field gradient measures the change of the electric field at the nucleus generated by the electronic charge distribution. We will see later that this can be described as a coupling between the two parameters (here tensors of rank 2) and this will give an energy splitting (in the order of μeV).

The magnetic hyperfine splitting has to do with the interaction of the nucleur dipole moment and the magnetic hyperfine field. We have seen such a magnetic field before, it is the magnetic field associated with the total angular momentum J. The coupling betwheen these two parameters (both vectors) will give our energy splitting (again in the order of μeV).

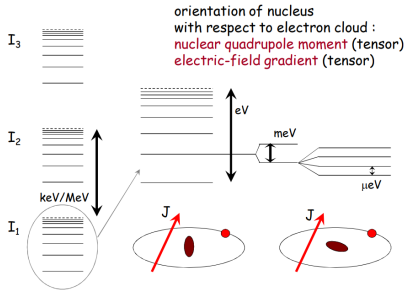

Figure 3: Illustration of the different terms in the electric hyperfine splitting.

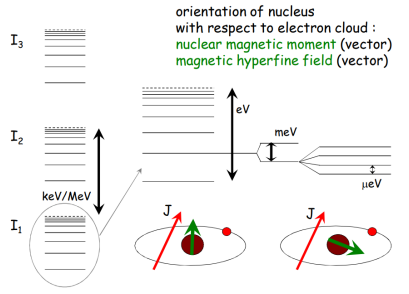

Figure 4: Illustration of the different terms in the magnetic hyperfine splitting.

2 Gravitational Analogue

This section extensively discusses a problem from classical mechanics. We will highly benefit from this when discussing the quantum multipole expansion and the quadrupole term where very similar reasonings appear. The full mathematics will be developed only once here, where we can profit from the absence of quantum mechanics, which could possibly distract us. In the later parts about the multipole expansion and the quadrupole term we will just have to copy the results, and concentrate on the interpretation. Sometimes notation becomes weird in this chapter. We prefer however to give the mathematical objects very explicit names, to point out the often subtle differences between them and hence avoid misconceptions.

2.1 Two Mass Distributions: Multipole Expansion

Take a look at the image bellow. How can we find the potential energy E_{pot} of said static system of two bodies M_1 and M_2 in situation 1? Where both masses are in each others gravitational fields, with their center of mass separated by a vector $\vec{r_0}$.

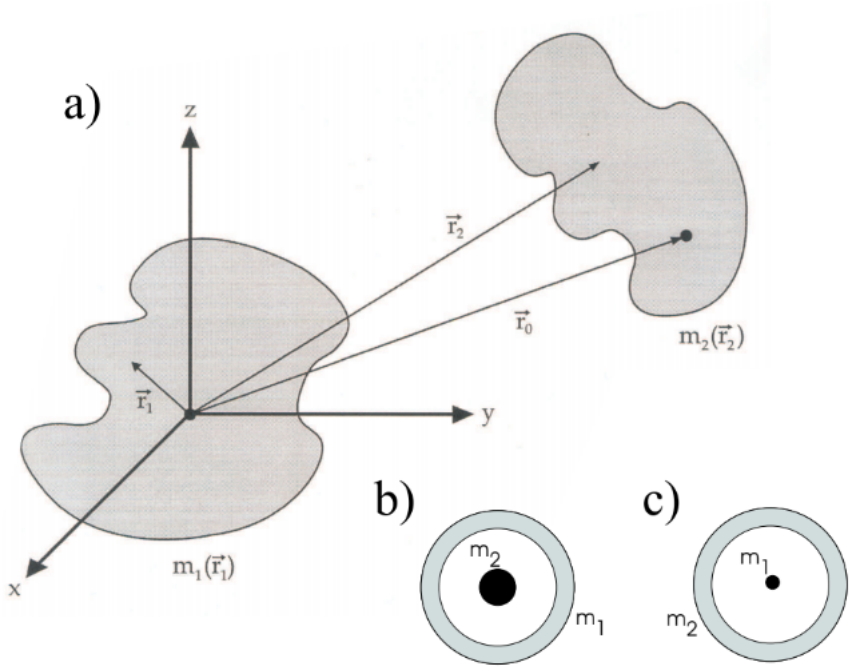

Figure 5: Three situations of two mass distributions interacting via gravitation.

We use the following notation: M_i is the name of the body, m_i its total mass, and $\rho_i(\vec{r_i})$ is its mass density distribution function. Our reference frame XYZ will have its centre in the barycentre of M_1[3]. In general, both mass distributions are inhomogeneous and have an irregular shape. Because potential energy is defined only apart from an additive constant, we make the usual convention that E_{pot} vanishes if the two masses are at an infinite distance from each other.

We can choose to calculate either the potential energy of M_2 in the field of M_1, or vice-versa. Choosing the latter possibility, we can write:

$$E_{pot} = \int_1 \rho_1(\vec{r}_1)\, V_2(\vec{r}_1)\, d\vec{r}_1 \tag{1}$$

The integral is taken over the volume occupied by M_1, or over all space. The potential V_2 of M_2 at $\vec{r}_1$ can be written as:

$$V_2(\vec{r}_1) = -G \int_2 \frac{\rho_2(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|}\, d\vec{r}_2 \tag{2}$$

which leads to the following expression for the potential energy:

$$E_{pot} \quad = \quad -G \int_1 \int_2 \frac{\rho_1(\vec{r}_1)\rho_2(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|}\, d\vec{r}_1\, d\vec{r}_2 \tag{3}$$

Due to the possibly irregular shapes (as will be in most general cases) of both mass distributions, the integrals in equations 1 to 3 can be hard to calculate. In order to be able to deal with simpler integrals and in order to gain simultaneously physical insight, we will make a series expansion of equation 3 using the so-called *Laplace expansion* or *multipole expansion* in spherical coordinates[4]

$$\frac{1}{|\vec{r}_2 - \vec{r}_1|} = 4\pi \sum_{n,q} \frac{r_<^n}{r_>^{n+1}} \frac{1}{2n+1}\, Y_q^{n*}(\theta_1,\, \phi_1) Y_q^n(\theta_2,\, \phi_2) \tag{4}$$

with $r_< = min\,(r_1,\, r_2)$ and $r_> = max\,(r_1,\, r_2)$. The potential energy of equation 3 then becomes:

$$E_{pot} = -4\pi\, G \int_1 \int_2 \rho_1(\vec{r}_1)\rho_2(\vec{r}_2) \left(\sum_{n,q} \frac{r_<^n}{r_>^{n+1}} \frac{1}{2n+1}\, Y_q^{n*}(\theta_1,\, \phi_1) Y_q^n(\theta_2,\, \phi_2) \right) d\vec{r}_1\, d\vec{r}_2 \tag{5}$$

[3] Any other centre will yield the same mathematics, but using this condition makes further calculations a lot easier.

[4] The Condon-Shortley phase convention for the spherical harmonics is used, see http://mathworld.wolfram.com/Condon-ShortleyPhase.html

In general, this is still a sum of very complicated integrals, as we cannot separate the integration over $\vec{r}_1$ and $\vec{r}_2$. However, if the two bodies are such that **any** r_1 **is smaller then any** r_2 [5], the separation can be made and we obtain:

$$E_{pot} = \sum_{n,q} Q_q^{n*} V_q^n \tag{6}$$

with

$$Q_q^n = \sqrt{\frac{4\pi}{2n+1}} \int_1 \rho_1(\vec{r}_1)\, r_1^n\, Y_q^n(\theta_1,\, \phi_1)\, d\vec{r}_1 \tag{7}$$

and

$$V_q^n = -G\sqrt{\frac{4\pi}{2n+1}} \int_2 \frac{\rho_2(\vec{r}_2)}{r_2^{n+1}}\, Y_q^n(\theta_2,\, \phi_2)\, d\vec{r}_2 \tag{8}$$

The Q-tensors have units $\mathrm{kg\,m}^n$, the V-tensors $\mathrm{N}/(\mathrm{kg\,m}^{n-1})$ and their products Nm or J. It is important to realize that the summation is a dot product between two spherical tensors (a different prefactor can occur if the dot product is taken between cartesian tensors).

2.2 The monopole term (n=0)

The monopole term can be read as a dot product between two tensors of rank 0 (scalars): the monopole moment Q_0^0 due to M_1 (units: kg), and the monopole field V_0^0 due to M_2 (units: Nm/kg). Explicit expressions are:

$$Q_0^0 = m_1 \tag{9}$$

$$V_0^0 = -G \int \frac{\rho_2(\vec{r}_2)}{|\vec{r}_2|}\, d\vec{r}_2 \tag{10}$$

$$E_{pot}^{(0)} = Q_0^{0*} V_0^0 \tag{11}$$

The monopole field is nothing else than the gravitational potential at the origin (where the barycentre of M_1 is) due to M_2, while the monopole moment is the total mass m_1 of M_1. The monopole contribution to the potential energy would be the only and exact contribution to the potential energy in the case where M_1 would be a point mass, situated at the origin.

[5] This excludes *a)* bodies that overlap (not possible for masses, but possible for charges: e.g. s-electron penetration in the nucleus), and *b)* a body with a hole in which a bulge on the other body enters.

2.3 The dipole term (n=1)

The dipole term can be read as a dot product between two tensors of rank 1 (vectors): the dipole moment Q_q^1 due to M_1 (units: kg m), and the dipole field V_q^1 due to M_2 (units: N/kg). Explicit expressions are:

$$Q_q^1 = \sqrt{\frac{4\pi}{3}} \int_1 \rho_1(\vec{r}_1)\, r_1\, Y_q^1(\theta_1,\, \phi_1)\, d\vec{r}_1 \tag{12}$$

$$V_q^1 = -G\sqrt{\frac{4\pi}{3}} \int_2 \frac{\rho_2(\vec{r}_2)}{r_2^2}\, Y_q^1(\theta_2,\, \phi_2)\, d\vec{r}_2 \tag{13}$$

$$E_{pot}^{(1)} = \sum_{q=-1,0,1} Q_q^{1*} V_q^1 \tag{14}$$

We can transform the dipole moment into 3 components of a cartesian vector, which will be more easily interpretable:

$$Q_x = \frac{\sqrt{2}}{2}\left(Q_{-1}^1 - Q_{+1}^1\right) \tag{15}$$

$$= \int_1 \rho_1(\vec{r}_1)\, r_1\, \sin\theta \cos\phi\, d\vec{r}_1 \tag{16}$$

$$= \int_1 \rho_1(\vec{r}_1)\, x_1\, d\vec{r}_1 \tag{17}$$

$$Q_y = \int_1 \rho_1(\vec{r}_1)\, y_1\, d\vec{r}_1 \tag{18}$$

$$Q_z = \int_1 \rho_1(\vec{r}_1)\, z_1\, d\vec{r}_1 \tag{19}$$

One recognizes the definition of the position vector of the center of mass of M_1, multiplied by the total mass m_1. As we have chosen the origin of the axis system in the center of mass, we can conclude that the three components of the dipole moment are zero, both in the cartesian and in the spherical form.

In a similar way, the following cartesian components are found for the dipole field:

$$V_x = -G \int_2 \frac{\rho_2(\vec{r}_2)}{|r_2|^3}\, x_2\, d\vec{r}_2 \tag{20}$$

$$V_y = -G \int_2 \frac{\rho_2(\vec{r}_2)}{|r_2|^3}\, y_2\, d\vec{r}_2 \tag{21}$$

$$V_z = -G \int_2 \frac{\rho_2(\vec{r}_2)}{|r_2|^3}\, z_2\, d\vec{r}_2 \tag{22}$$

One recognizes now that the dipole field vector is the opposite of the gravitational field due to M_2 at the origin[6].

2.4 The quadrupole term (n=2)

The quadrupole term can be read as a dot product between two tensors of rank 2: the quadrupole moment Q_q^2 due to M_1 (units: $\mathrm{kg\,m^2}$), and the quadrupole field V_q^2 due to M_2 (units: $\mathrm{N/(kg\,m)}$). Explicit expressions are:

$$Q_q^2 = \sqrt{\frac{4\pi}{5}} \int_1 \rho_1(\vec{r}_1)\, r_1^2\, Y_q^2(\theta_1,\, \phi_1)\, d\vec{r}_1 \tag{24}$$

$$V_q^2 = -G\sqrt{\frac{4\pi}{5}} \int_2 \frac{\rho_2(\vec{r}_2)}{r_2^3}\, Y_q^2(\theta_2,\, \phi_2)\, d\vec{r}_2 \tag{25}$$

$$E_{pot}^{(2)} = \sum_{q=-2,\ldots,2} Q_q^{2*} V_q^2 \tag{26}$$

Being given a spherical tensor field of rank 2, the corresponding 6 components of its cartesian form (only 5 of them are independent) are found by:

$$\begin{aligned} a_{11} &= \frac{\sqrt{6}}{2}\left(a_2^2 + a_{-2}^2\right) - a_0^2 \\ a_{22} &= -\frac{\sqrt{6}}{2}\left(a_2^2 + a_{-2}^2\right) - a_0^2 \\ a_{33} &= 2a_0^2 \\ a_{12} &= -\frac{\sqrt{6}}{2} i \left(a_2^2 - a_{-2}^2\right) \\ a_{13} &= -\frac{\sqrt{6}}{2}\left(a_1^2 - a_{-1}^2\right) \\ a_{23} &= \frac{\sqrt{6}}{2} i \left(a_1^2 + a_{-1}^2\right) \end{aligned} \tag{27, 28}$$

The quadrupole moment tensor can be transformed in its cartesian form: a traceless, symmetric matrix:

$${}_cQ_{sh}^{(2)} = \int_1 \rho_1(\vec{r}_1) \begin{bmatrix} 3x_1^2 - r_1^2 & 3x_1y_1 & 3x_1z_1 \\ 3x_1y_1 & 3y_1^2 - r_1^2 & 3y_1z_1 \\ 3x_1z_1 & 3y_1z_1 & 3z_1^2 - r_1^2 \end{bmatrix} d\vec{r}_1 \tag{29}$$

[6] $\vec{E}_2(\vec{0}) = -\vec{\nabla} V_2(\vec{0})$, with V_2 given by equation 2. This results in:

$$\vec{E}_2(\vec{0}) = G \int_2 \frac{\rho_2(\vec{r}_2)}{|r_2|^3}\, \vec{r}_2\, d\vec{r}_2 \tag{23}$$

It is understood that the integration over $\vec{r}_1$ is performed for all of the 9 elements. The physical interpretation of the quadrupole moment tensor is as follows: its i^{th} diagonal element will be positive if along the i^{th} axis of the reference frame the actual radius of M_1 is larger than the radius of the best-approximating sphere. In this situation M_1 is said to be *prolate* along this axis. In the inverse case, M_1 is *oblate* along this axis (see example in Fig. 6). For a perfect sphere, this tensor is zero.

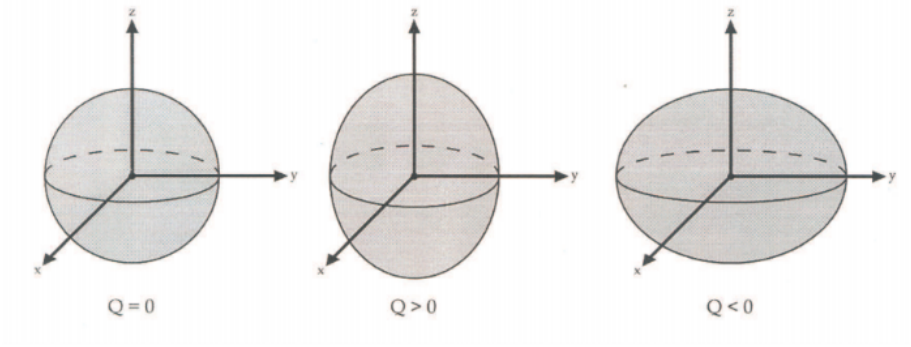

Figure 6: *A spherical, prolate and oblate mass distribution (with respect to the z-axis).*

The same transformation (using equation 27) can be done for the quadrupole field:

$$ {}_cV_{sh}^{(2)} = -G \int_2 \frac{\rho_2(\vec{r}_2)}{|\vec{r}_2|^5} \begin{bmatrix} 3x_2^2 - r_2^2 & 3x_2y_2 & 3x_2z_2 \\ 3x_2y_2 & 3y_2^2 - r_2^2 & 3y_2z_2 \\ 3x_2z_2 & 3y_2z_2 & 3z_2^2 - r_2^2 \end{bmatrix} d\vec{r}_2 \quad (30) $$

Note that the structure of this quadrupole field tensor is definitely different from the quadrupole moment tensor, due to the factor $1/|\vec{r}_2|^5$.

How to interpret the meaning of this tensor? Let us take the negative gradient of the x-component of the gravitational field:

$$ -\nabla E_{2x}(\vec{0}) = \left(-\frac{\partial E_{2x}(\vec{0})}{\partial x_1}, -\frac{\partial E_{2x}(\vec{0})}{\partial y_1}, -\frac{\partial E_{2x}(\vec{0})}{\partial z_1} \right) \quad (31) $$

$$ = \left(\frac{\partial^2 V_2(\vec{0})}{\partial x_1^2}, \frac{\partial^2 V_2(\vec{0})}{\partial y_1 \partial x_1}, \frac{\partial^2 V_2(\vec{0})}{\partial z_1 \partial x_1} \right) \quad (32) $$

Repeating this for the y- and z-components of the gravitational field leads to 9 quantities which can be arranged in a symmetric matrix:

$$\begin{bmatrix} \frac{\partial^2 V_2(\vec{0})}{\partial x_1^2} & \frac{\partial^2 V_2(\vec{0})}{\partial y_1 \partial x_1} & \frac{\partial^2 V_2(\vec{0})}{\partial z_1 \partial x_1} \\ \frac{\partial^2 V_2(\vec{0})}{\partial x_1 \partial y_1} & \frac{\partial^2 V_2(\vec{0})}{\partial y_1^2} & \frac{\partial^2 V_2(\vec{0})}{\partial z_1 \partial y_1} \\ \frac{\partial^2 V_2(\vec{0})}{\partial x_1 \partial z_1} & \frac{\partial^2 f(\vec{0})}{\partial y_1 \partial z_1} & \frac{\partial^2 V_2(\vec{0})}{\partial z_1^2} \end{bmatrix} \tag{33}$$

This matrix is also trace-less. Indeed, its trace is the Laplacian of the potential due to M_2, evaluated at the origin:

$$\Delta V_2(\vec{0}) \;=\; \frac{\partial^2 V_2(\vec{0})}{\partial x^2} + \frac{\partial^2 V_2(\vec{0})}{\partial y^2} + \frac{\partial^2 V_2(\vec{0})}{\partial z^2} \;=\; 4\pi\, G\, \rho_2(\vec{0}) \tag{34}$$

By the Poisson equation, this Laplacian can be related to $\rho_2(\vec{0})$, the mass density of M_2 at the origin. By our restriction that $\vec{r}_1 < \vec{r}_2$, $\rho_2(\vec{0})$ must necessarily be zero, and the above matrix is trace-less. You can verify now that:

$$V_{ij} \;=\; \frac{\partial^2 V_2(\vec{0})}{\partial x_{1i} \partial x_{1j}} \tag{35}$$

$$\;=\; -G \int_2 \frac{\rho_2(\vec{r}_2)}{|\vec{r}_2|^5} \left(3 x_{2i} x_{2j} - r_2^2 \delta_{ij}\right) d\vec{r}_2 \tag{36}$$

which are exactly the 9 components of the cartesian form of the quadrupole field given in equation 30. Looking at equations 31 and 33, we can therefore interpret the quadrupole field tensor as the negative gradient of the gravitational field at the origin due to M_2. Therefore the quadrupole field tensor is often called the *(gravitational-) field gradient* tensor. Its ij^{th} element expresses how strongly the i-component of the gravitational field at the origin varies if one goes along the j-direction.

2.5 Correction to Multipole Expansion.

By assuming $r_1 < r_2$ in the Laplace expansion, we made an error for those mass distributions where this condition is not fulfilled. Often this error will be small, but in the case of overlapping charge distributions as we will meet them in the following chapters, it will produce nevertheless measurable effects. The necessary corrections can be expressed as corrections to each of the multipole terms separately, but the correction to the monopole term is the most important one. In order to find this correction, we will take the opportunity to use the multipole expansion in cartesian coordinates rather than the Laplace expansion, which sheds a different light on the problem we have just solved in spherical coordinates. The full solution using spherical coordinates involves Tesseral harmonics, and can be found at many places.

We start again from equation 3. In order to avoid the complicated integral in this expression, we make a Taylor expansion of V_2 around the origin of the axis system. Such a Taylor expansion converges rapidly if the points $\vec{r}_1$ at which we need the value of V_2 are much closer to the origin than the major part of the mass of M_2 is. We assume that this is the case[7], and will truncate the series after the second order term. On how to expand a potential, consider a function

$$f(\vec{r}) \quad = \quad \int \frac{g(\vec{r}_v)}{|\vec{r}_v - \vec{r}|}\, d\vec{r}_v \tag{37}$$

The integral runs over that part of space where $g(\vec{r}_v)$ is not zero, which might be a finite or infinite region. If g is a charge or mass distribution, f gives the electric or gravitational potential in a point $\vec{r}$ (apart from an appropriate factor). That point can be either inside or outside the non-zero region of g. If it lies inside, the denominator in the integral becomes zero and we have to care about the convergence of the integral. The latter is determined by the properties of g. We assume that we know the value of f and of all its derivatives *at the origin* $\vec{0}$. What we want to know is the value of f at points $\vec{r} = (x, y, z)$ that are not far away from $\vec{0}$. This means we need a Taylor expansion of $f(\vec{r})$ around $\vec{0}$. The general form of a Taylor expansion around $\vec{0}$ for a function with vectors as argument, is:

$$f(\vec{0} + \vec{r}) \quad = \quad \sum_{j=0}^{\infty} \left[\frac{1}{j!} \left(\vec{r} \cdot \vec{\nabla}_{\vec{r}\,'} \right)^j \, f(\vec{r}\,') \right]_{\vec{r}\,'=\vec{0}} \tag{38}$$

Explicitly for our case, this gives for the zeroth order term:

$$\begin{aligned} E_{pot}^{(0)} &= \left(\int \rho_1(\vec{r}_1)\, d\vec{r}_1 \right) V_2(\vec{0}) && (39)\\ &= m_1\, V_2(\vec{0}) && (40)\\ &= m_1 \left(-G \int \frac{\rho_2(\vec{r}_2)}{|\vec{r}_2|}\, d\vec{r}_2 \right) && (41)\\ &= Q_0^0 V_0^0 && (42)\\ &= {}_sQ_{sh}^{(0)} \cdot {}_sV_{sh}^{(0)} \;=\; {}_cQ_{sh}^{(0)} \cdot {}_cV_{sh}^{(0)} && (43) \end{aligned}$$

In equation 41, we recognize the monopole moment and monopole field derived in equations 9 and 10.

[7] We will use this later for atoms, where $\vec{r}_1$ is of the order of the nuclear radius (10^{-15} m) and $\vec{r}_2$ of the order of the radius of an electron orbit (10^{-10} m).

The first order term in the expansion of E_{pot} can be written as:

$$E^{(1)}_{pot} = \left[\int \rho_1(\vec{r}_1) x_1 \, d\vec{r}_1 \quad \int \rho_1(\vec{r}_1) y_1 \, d\vec{r}_1 \quad \int \rho_1(\vec{r}_1) z_1 \, d\vec{r}_1 \right] \begin{bmatrix} \frac{\partial V_2(\vec{0})}{\partial x_1} \\ \frac{\partial V_2(\vec{0})}{\partial y_1} \\ \frac{\partial V_2(\vec{0})}{\partial z_1} \end{bmatrix} \tag{44}$$

$$= \left[\int \rho_1(\vec{r}_1) x_1 \, d\vec{r}_1 \quad \int \rho_1(\vec{r}_1) y_1 \, d\vec{r}_1 \quad \int \rho_1(\vec{r}_1) z_1 \, d\vec{r}_1 \right] \begin{bmatrix} -G \int \frac{\rho_2(\vec{r}_2)}{|\vec{r}_2|^3} x_2 \, d\vec{r}_2 \\ -G \int \frac{\rho_2(\vec{r}_2)}{|\vec{r}_2|^3} y_2 \, d\vec{r}_2 \\ -G \int \frac{\rho_2(\vec{r}_2)}{|\vec{r}_2|^3} z_2 \, d\vec{r}_2 \end{bmatrix} \tag{45}$$

$$= {}_cQ^{(1)}_{sh} \cdot {}_cV^{(1)}_{sh} = {}_sQ^{(1)}_{sh} \cdot {}_sV^{(1)}_{sh} \tag{46}$$

We recognize the cartesian forms of the dipole moment and the dipole field, as derived before.
The second order term in the expansion of E_{pot} is:

$$E^{(2)}_{pot} = \frac{1}{2} \begin{bmatrix} \int \rho_1(\vec{r}_1) x_1^2 \, d\vec{r}_1 & \int \rho_1(\vec{r}_1) x_1 y_1 \, d\vec{r}_1 & \int \rho_1(\vec{r}_1) x_1 z_1 \, d\vec{r}_1 \\ \int \rho_1(\vec{r}_1) y_1 x_1 \, d\vec{r}_1 & \int \rho_1(\vec{r}_1) y_1^2 \, d\vec{r}_1 & \int \rho_1(\vec{r}_1) y_1 z_1 \, d\vec{r}_1 \\ \int \rho_1(\vec{r}_1) z_1 x_1 \, d\vec{r}_1 & \int \rho_1(\vec{r}_1) z_1 y_1 \, d\vec{r}_1 & \int \rho_1(\vec{r}_1) z_1^2 \, d\vec{r}_1 \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial^2 V_2(\vec{0})}{\partial x_1^2} & \frac{\partial^2 V_2(\vec{0})}{\partial y_1 \partial x_1} & \frac{\partial^2 V_2(\vec{0})}{\partial z_1 \partial x_1} \\ \frac{\partial^2 V_2(\vec{0})}{\partial x_1 \partial y_1} & \frac{\partial^2 V_2(\vec{0})}{\partial y_1^2} & \frac{\partial^2 V_2(\vec{0})}{\partial z_1 \partial y_1} \\ \frac{\partial^2 V_2(\vec{0})}{\partial x_1 \partial z_1} & \frac{\partial^2 f(\vec{0})}{\partial y_1 \partial z_1} & \frac{\partial^2 V_2(\vec{0})}{\partial z_1^2} \end{bmatrix} \tag{47}$$

$$= \frac{1}{2} \, {}_cK^{(2)} \cdot {}_cW^{(2)} \tag{48}$$

This is a dot product between two (Cartesian) tensors of rank 2 (mind the fact that this is no matrix multiplication, but short-hand notation for a dot product)[8]. In contrast to the zeroth and first order terms, we cannot immediately identify the two cartesian tensors in this dot product with multipole moments and multipole fields. The left tensor seems to be related to the quadrupole moment (equation 29), but is not identical to it. The right tensor seems at first

[8] We are all accustomed to the dot product for 2 vectors of dimension n. Where we multiply the first element of the first vector with the first element of the second vector and sum it with the product between the second element of the first vector and the second element of the second vector and so on to the n'th pair of elements. It is clear that both vectors need to be of the same dimention. The same holds for the dot product between two matrices (dimensions n x m). Here we will multiply element (0,0) of the first matrix with element (0,0) of the second matrix and sum it with the product between (0,1) of the first matrix and element (0,1) of the second matrix, all the way to (0,m). Then we continue summing for (1,0) to (1,m). Up to (n,0) to (n,m). Summing each product until we get our final result (a scalar).

sight identical to the cartesian form of the quadrupole field (equation 30), but it is not: its trace – the Laplacian of the potential – is not necessarily zero, because we did not need to require $r_1 < r_2$ for the Taylor expansion. By equation 34, this trace can be different from zero. But we can write in the same way both tensors as a sum of the quadrupole moment/field and a multiple of the unit 3×3 matrix (the integrations are noted in short-hand by curled {brackets}):

$$ {}_cK^{(2)} = \frac{1}{3}\begin{bmatrix} \{3x_1^2\} - \{r_1^2\} & \{3x_1y_1\} & \{3x_1z_1\} \\ \{3y_1x_1\} & \{3y_1^2\} - \{r_1^2\} & \{3y_1z_1\} \\ \{3z_1x_1\} & \{3z_1y_1\} & \{3z_1^2\} - \{r_1^2\} \end{bmatrix} + \frac{1}{3}\begin{bmatrix} \{r_1^2\} & 0 & 0 \\ 0 & \{r_1^2\} & 0 \\ 0 & 0 & \{r_1^2\} \end{bmatrix} \tag{49} $$

$$ {}_cW^{(2)} = \begin{bmatrix} \frac{\partial^2 V_2(\vec{0})}{\partial x_1^2} - \frac{\Delta V_2(\vec{0})}{3} & \frac{\partial^2 V_2(\vec{0})}{\partial y_1 \partial x_1} & \frac{\partial^2 V_2(\vec{0})}{\partial z_1 \partial x_1} \\ \frac{\partial^2 V_2(\vec{0})}{\partial x_1 \partial y_1} & \frac{\partial^2 V_2(\vec{0})}{\partial y_1^2} - \frac{\Delta V_2(\vec{0})}{3} & \frac{\partial^2 V_2(\vec{0})}{\partial z_1 \partial y_1} \\ \frac{\partial^2 V_2(\vec{0})}{\partial x_1 \partial z_1} & \frac{\partial^2 f(\vec{0})}{\partial y_1 \partial z_1} & \frac{\partial^2 V_2(\vec{0})}{\partial z_1^2} - \frac{\Delta V_2(\vec{0})}{3} \end{bmatrix} + \begin{bmatrix} \frac{\Delta V_2(\vec{0})}{3} & 0 & 0 \\ 0 & \frac{\Delta V_2(\vec{0})}{3} & 0 \\ 0 & 0 & \frac{\Delta V_2(\vec{0})}{3} \end{bmatrix} \tag{50} $$

Now make the dot product between these two sums. You end up with 4 terms, of which two will be zero: the dot product between a trace-less tensor and a multiple of a unit matrix is zero. With the two nonzero terms we can write the second order term of the Taylor expansion of E_{pot} as:

$$ \begin{aligned} E_{pot}^{(2)} = {} & \frac{1}{6}\begin{bmatrix} \{3x_1^2\} - \{r_1^2\} & \{3x_1y_1\} & \{3x_1z_1\} \\ \{3y_1x_1\} & \{3y_1^2\} - \{r_1^2\} & \{3y_1z_1\} \\ \{3z_1x_1\} & \{3z_1y_1\} & \{3z_1^2\} - \{r_1^2\} \end{bmatrix} \cdot \\ & \begin{bmatrix} \frac{\partial^2 V_2(\vec{0})}{\partial x_1^2} - \frac{\Delta V_2(\vec{0})}{3} & \frac{\partial^2 V_2(\vec{0})}{\partial y_1 \partial x_1} & \frac{\partial^2 V_2(\vec{0})}{\partial z_1 \partial x_1} \\ \frac{\partial^2 V_2(\vec{0})}{\partial x_1 \partial y_1} & \frac{\partial^2 V_2(\vec{0})}{\partial y_1^2} - \frac{\Delta V_2(\vec{0})}{3} & \frac{\partial^2 V_2(\vec{0})}{\partial z_1 \partial y_1} \\ \frac{\partial^2 V_2(\vec{0})}{\partial x_1 \partial z_1} & \frac{\partial^2 f(\vec{0})}{\partial y_1 \partial z_1} & \frac{\partial^2 V_2(\vec{0})}{\partial z_1^2} - \frac{\Delta V_2(\vec{0})}{3} \end{bmatrix} + \\ & \frac{1}{6}\begin{bmatrix} \{r_1^2\} & 0 & 0 \\ 0 & \{r_1^2\} & 0 \\ 0 & 0 & \{r_1^2\} \end{bmatrix} \cdot \begin{bmatrix} \frac{\Delta V_2(\vec{0})}{3} & 0 & 0 \\ 0 & \frac{\Delta V_2(\vec{0})}{3} & 0 \\ 0 & 0 & \frac{\Delta V_2(\vec{0})}{3} \end{bmatrix} \end{aligned} \tag{51} $$

Working it out shows that this is just a scalar product between two numbers, which we can interpret as a dot product:

$$ \frac{1}{6} {}_cQ_{sz}^{(0)} \cdot {}_cV_{sz}^{(0)} = \frac{1}{6}\Delta V_2(\vec{0}) \left\langle r_1^2 \right\rangle \tag{52} $$

$$ = \frac{4\pi G}{6}\rho_2(\vec{0}) \int \rho_1(\vec{r}_1)\, r_1^2 \, d\vec{r}_1 \tag{53} $$

This energy term is the leading correction to the multipole expansion of equation 6 for situations where r_1 can be larger than r_2. It is a dot product between two tensors of rank 0, hence this is a correction to the monopole term of the multipole expansion. The reason why we find this contribution in the Taylor expansion and not in the (approximated) Laplace expansion is that in the former we did not require $r_1 < r_2$.

3 The Double Ring

Consider a dumb-bell consisting out of two equal point masses connected by a rigid massless rod. The total mass of the dumb-bell is m_1, the length of the rod is l_1. As the dumb-bell will play the role of M_1, we fix it with its centre of mass (the middle of the rod) to the origin of an axis system XYZ, in such a way that the dumb-bell can rotate freely about this origin. As second mass distribution take 2 rings with radius R in planes parallel to XY, separated by a distance h, one $\frac{h}{2}$ above the XY-plane and one $\frac{h}{2}$ below it. The total mass of the 2 rings is m_2, and its (constant) linear mass density is ρ_2 (fig. reffig-2-3). The question we want to solve is: if also the rings are kept fixed in XYZ, what will be the preferred (= lowest-energy) orientation of the dumb-bell?

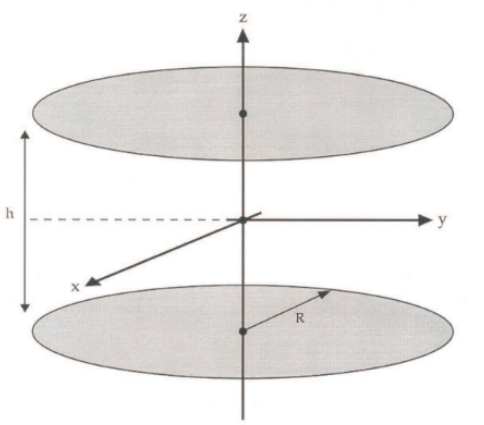

Figure 7: *A double ring system. The rings are shaded gray for clarity, but actually they are hollow and all mass is concentrated at their circumference.*

The condition $\frac{r_1}{r_2} \ll 1$ becomes here $\frac{l_1}{2} \ll \sqrt{\frac{h^2}{4} + R^2}$, which is fulfilled if the length of the dumb-bell is small enough compared to the geometry of the double ring. The problem has some circular symmetry, and it will be useful to use spherical coordinates in the cartesian tensors[9]. There is no overlap and hence

[9] An alternative is to use equation 5 up to the quadrupole term. Try this solution yourself.

no size-dependent monopole term. The shape-dependent monopole term can easily be calculated to be[10]:

$$\begin{aligned} E_{pot}^{(0)} &= V_{sh}^{(0)} \cdot Q_{sh}^{(0)} \\ &= -\frac{Gm_1m_2}{\sqrt{\frac{h^2}{4}+R^2}} \end{aligned} \tag{54}$$

For an arbitrary orientation (θ, ϕ) of the dumb-bell, its quadrupole moment is:

$${}_cQ_{sh}^{(2)} = \frac{3m_1l_1^2}{4} \begin{bmatrix} \sin^2\theta\cos^2\phi - \frac{1}{3} & \sin^2\theta\sin\phi\cos\phi & \sin\theta\cos\theta\cos\phi \\ \sin^2\theta\sin\phi\cos\phi & \sin^2\theta\sin^2\phi - \frac{1}{3} & \sin\theta\cos\theta\sin\phi \\ \sin\theta\cos\theta\cos\phi & \sin\theta\cos\theta\sin\phi & \cos^2\theta - \frac{1}{3} \end{bmatrix} \tag{55}$$

Note that this quadrupole moment does not change upon inversion of the axis system ($\theta \to \pi - \theta$, $\phi \to \phi + \pi$), which we expect as also the dumb-bell has inversion symmetry.

In order to calculate the tensor of the gradient of the gravitational field generated by the rings at the origin of XYZ, we use the following equalities:

$$\begin{cases} R = r_2 \sin\theta_0 \\ h = 2r_2\cos\theta_0 \\ \rho_2 = \frac{m_2}{4\pi\, r_2 \sin\theta_0} \\ R^2 + \frac{h^2}{4} = r_2^2 \end{cases} \tag{56}$$

and find that[11] (there is no overlap, such that $\rho_2(\vec{0}) = 0$):

$${}_cV_{sh}^{(2)} = -\frac{Gm_2\,(h^2-2R^2)}{8R(R^2+\frac{h^2}{4})^{\frac{5}{2}}} \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \tag{57}$$

The diagonal form of this equation shows that XYZ by chance (?) is a principal axis system for the tensor $V_{sh}^{(2)}$. After some straightforward goniometric manipulation, the quadrupole energy becomes:

$$\frac{1}{6}Q_{sh}^{(2)} \cdot V_{sh}^{(2)} = -\frac{3Gm_1m_2\,l_1^2\,(h^2-2R^2)}{32R(R^2+\frac{h^2}{4})^{\frac{5}{2}}}\,(2\cos^2\theta - \sin^2\theta) \tag{58}$$

This energy does not depend on the azimuthal orientation ϕ of the dumb-bell. The value of θ_{low} where the quadrupole energy is minimal will depend on the sign of $h^2 - 2R^2$:

$$\begin{cases} \sqrt{2}\,R < h & \Longrightarrow \quad \theta_{low} = 0^\circ \\ \sqrt{2}\,R = h & \Longrightarrow \quad \theta_{low} = \textit{any angle} \\ \sqrt{2}\,R > h & \Longrightarrow \quad \theta_{low} = 90^\circ \end{cases} \tag{59}$$

[10] Both centers of mass coincide at $\vec{0}$, hence $\vec{r}_0 = \vec{0}$ and one could write everything explicitly in terms of fields as $E_{pot}^{(0)}(\vec{0})$ etc.

[11] take $d\vec{r}_2 = 2\,r_2 \sin\theta_0\, d\phi$ (the factor 2 is due to the *double* ring)

If the radius is small enough compared to the distance between the rings, the dumb-bell has it lowest energy when it lies parallel to the Z-axis. Not unexpected: the limiting case are two rings which are so small that they are almost point masses on the z-axis, and obviously the dumb-bell would like to lie along the z-axis in such case, as this minimizes the distance between the dumb-bell and the masses, and hence maximizes the gravitational attraction. When the radius is large, the dumb-bell prefers an orientation in the XY-plane. Also not unexpected: the limiting case is two coinciding rings, and now distances are minimized if the dumb-bell is in the plane of the rings. One special R/h-ratio exists for which the orientation of the dumb-bell does not matter.

It is instructive to have a closer look at this. In section 2.4, we defined the *main component* of a rank 2 tensor, and mass distributions are divided into three distinct classes based on the sign of their main component (see Figure 6). With this terminology, we see that θ_{low} for the dumb-bell being 0°, anything or 90° corresponds to the main component of the double ring's field gradient being negative, zero or positive (equation 57). Three typical double ring systems belonging to each of the 3 classes are drawn below.

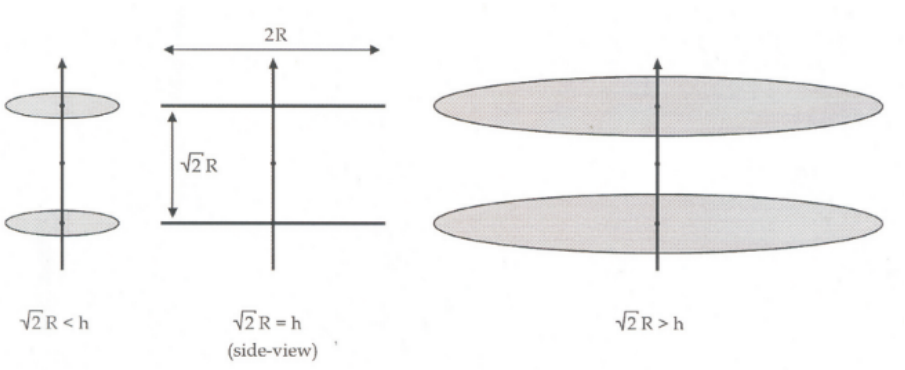

Figure 8: *The three distinct classes of double-ring systems.*

For some orientations the quadrupole correction is negative and for others positive. If the total energy is approximated by the sum of monopole and quadrupole contributions, some orientations will reduce the total energy and others will increase it. This is shown schematically in Fig. 9, a kind of picture we will meet again later in a quantummechanical situation. Note that all quadrupole energies between the two outer values are possible (i.e. all angles θ are possible).

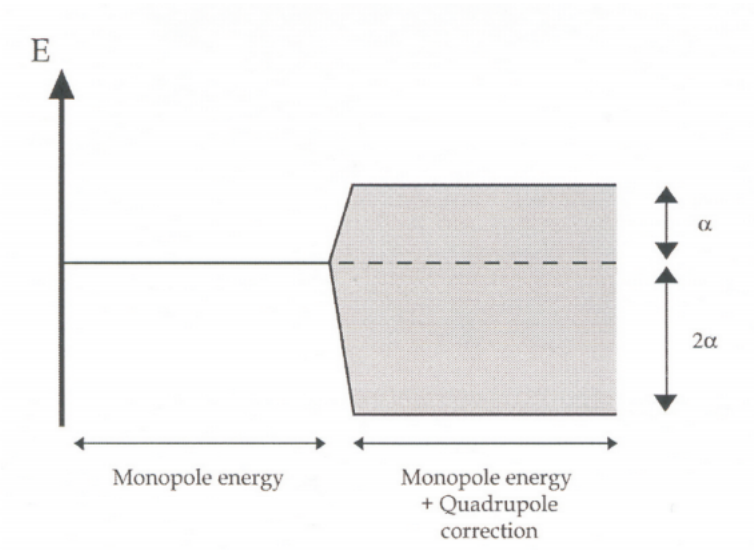

Figure 9: *Quadrupole corrections to the monopole energy for a dumb-bell in a double ring, for the case where* $h > \sqrt{2}R$. *For any orientation of the dumb-bell, a total energy between the shown outer values is found. The quantity* α *has the form* $\alpha = \frac{Gm_1m_2\ell_1^2(h^2-2R^2)}{32(R^2+\frac{h^2}{4})^{5/2}}$