Modern Physics for students of Software Engineering

Doron Cohen
Department of Physics, Ben-Gurion University, Beer-Sheva 84105, Israel

(based on arXiv:quant-ph/0605180)

This subset of lectures is based on lecture notes in quantum mechanics that are intended for Physics students. They were adjusted for a Modern Physics Course for students of Software Engineering.

Basic Physics 2

1 Basic concepts . . . 2
2 Semiclassical Dynamics . . . 10
3 Classical oscillations and waves . . . 17
4 Electricity . . . 23
5 Stochastic Dynamics . . . 29

Mathematics 33

6 Algebra . . . 33
7 Translations and momentum . . . 38
8 Rotations and Spin . . . 41
* Polarization of photons and electrons . . . [link]
* Quantum Encryption . . . [link]

Quantum Mechanics 48

9 The evolution of quantum mechanical systems . . . 48
10 Dynamics of an N site system . . . 53
11 Theory of quantum computation . . . 55
* optional source - Part I . . . [link]
* optional source - Part II . . . [link]

Fundamentals 67

12 The uncertainty principle . . . 67
* Realism and Determinism . . . [link]
* EPR and Bell inequality . . . [link]
* Quantum Measurements . . . [link]

Basic Physics

[1] Basic concepts

תמלול של הרצאת פתיחה בעברית בקישור הבא

Introduction to Modern Physics

[1.1] The building blocks of the universe [can be skipped]

The universe consists of a variety of particles which are described by the *standard model*. The known particles are divided into two groups:

- Quarks: constituents of the proton and the neutron, which form the ~ 100 nuclei known to us.
- Leptons: include the electrons, muons, taus, and the neutrinos.

The interaction between the particles is via fields (direct interaction between particles is contrary to the principles of the *special theory of relativity*). These interactions are responsible for the way material is organized. On top we have Gravitation (aka *general theory of relativity*).

In the present course we focus on the dynamics of a particle in a discrete space that consists of N sites. This simplified setting is enough in order to clarify the foundations of quantum mechanics. Contrary to the common pedagogical approach, the continuum limit (particle is 3D space) will not be discussed, and will not be taken as a starting point.

[1.2] Stochastic dynamics [can be skipped]

The simplest type of dynamics assumes that the "phase-space" of the particle consists of N locations (sites) labeled by a coordinate x_n, where $n = 1, ..., N$. The state of the particle in described by a probability function P_n, such that $\sum_n P_n = 1$. The *pure-state* of being with 100% probability in site n is denoted $|n\rangle$. The dynamics is described by a rate equation

$$\frac{d}{dt}\boldsymbol{P} = \mathcal{W}\boldsymbol{P} \tag{1.1}$$

with the solution

$$\boldsymbol{P}(t) = e^{\mathcal{W}t}\boldsymbol{P}(0) \tag{1.2}$$

where $\boldsymbol{P}$ is a column, and $\mathcal{W}$ is a matrix. The off-diagonal elements are the rates of transitions, namely, w_{nm} is the rate of transition from m to n. The diagonal elements $-\gamma_i$ of the $\mathcal{W}$ matrix are determined such that each column sums to zero. Accordingly we get conservation of probability, namely,

$$\frac{d}{dt}\sum_n P_n = 0 \tag{1.3}$$

Simple examples for stochastic dynamics will be introduced in the next section. In general the dynamics features a relaxation process that leads to a steady state. If the transitions are induced by a thermal environment, the

steady state is known as equilibrium state. If we add external fields the system may approach a non-equilibrium steady state (for example a ring with with a steady state current).

[1.3] Quantum dynamics [can be skipped]

The "phase-space" of the particle is an "Hilbert space". For a particle in N site system this Hilbert space is spanned by the states $|x_n\rangle$. For example $|\psi\rangle = \psi_1 |x_1\rangle + \psi_2 |x_2\rangle + \psi_3 |x_3\rangle$ is a possible physical *pure-state* of the system.

The (statistical) state of the particle in described by a probability matrix ρ. The diagonal elements of ρ are the probabilities p_n, and the off-diagonal elements are complex numbers, aka *coherences*. The probability matrix for a system that is *prepared* in a pure state $|\psi\rangle = \sum_n \psi_n |x_n\rangle$ is $\rho_{nm} = \psi_n \psi_m^*$. The diagonal elements are the probabilities $P_n = |\psi_n|^2$, while the off-diagonal elements are characterized by two phases φ_{nm}.

The postulates of quantum mechanics imply that the evolution of a *pure* state is dictated by an equation of the form

$$\frac{d}{dt}\psi = -i\mathcal{H}\psi \tag{1.4}$$

with the solution

$$\psi(t) = e^{-i\mathcal{H}t}\psi(0) \tag{1.5}$$

The matrix $\mathcal{H}$ is called Hamiltonian.

[1.4] Classical dynamics [can be skipped]

Classical dynamics can be regarded as some kind of simplified description of quantum dynamics. The *phase-space* of the particle consists of classical-like states $|\bar{x}, \bar{p}\rangle$. A classical-like state is called "wave-packet". It is characterized by an average position $\bar{x}$, and by a superposition phase that is defined per unit distance, namely $\bar{p} = \varphi/a$, where a is the distance between sites.

We shall see that classical dynamics can be generated by Hamiltonian function $H(x, p)$ such that

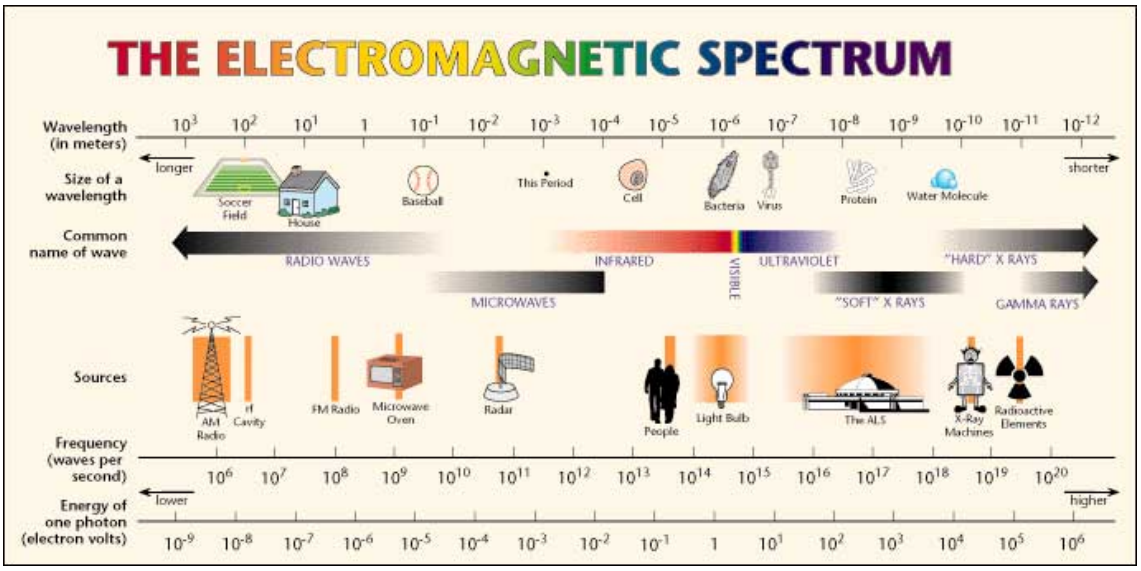
$$\dot{x} = \frac{\partial H}{\partial p} \tag{1.6}$$

$$\dot{p} = -\frac{\partial H}{\partial x} \tag{1.7}$$

We have used $|\bar{x}, \bar{p}\rangle$ to label classical-like states of the particle. Note that the x of the previous section was merely a site index, while $\bar{x}$ can be regarded as a dynamical variable that changes with time. The assumption of classical mechanics is that we start with a classical-lie state that is characterized by some initial values ($\bar{x}$,$\bar{p}$), and that the subsequent evolution does not spoil its classical-like feature. So at any later moment we have a classical-like state $|\bar{x}(t), \bar{p}(t)\rangle$, that can be visualized as a point (or better to say a small distribution of points) in phase-space (x, p).

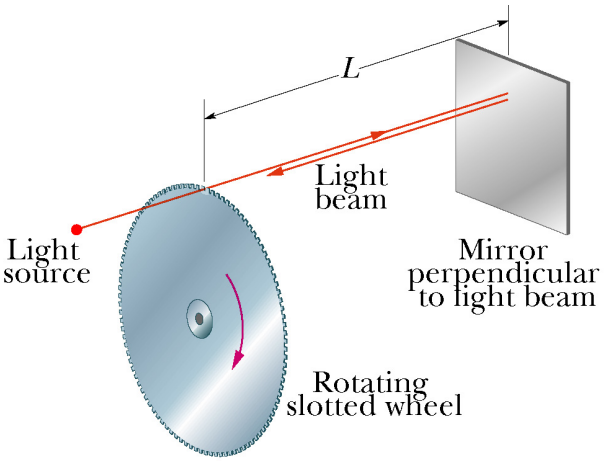
[1.5] Waves

A beam of electrons behaves as a "wave". Also a beam of light behaves as a "wave". The latter is a spacial case of electromagnetic wave (see figure below). In both cases we have to "filter" the wave in order to have a *plain wave* that has a well defined wavelength λ. We define for particles the momentum as $p = 2\pi/\lambda$, and for electromagnetic wave the wavenumber as $k = 2\pi/\lambda$. As a matter of style, depending on the context, we use the notation p or k for the same mathematical concept. The dispersion relation relates the velocity v of the beam to its wavenumber (k) or momentum (p). For electromagnetic wave in vacuum $v(k) = c$ does not depend on k. For massive particles $v(p) \approx (1/\mathsf{m})p$ depends on p. Below we discuss how velocity and wavelength are measured.



[1.6] Measurement of velocity

בשנת 1849 פיזו ביצע ניסוי למדידת מהירות האור. הוא שלח קרן אור שעברה בחריץ שבין שתי שיניים בגלגל מסתובב, והוחזרה על ידי מראות לאותו גלגל שיניים, לאחר שעברה מרחק מסוים. פיזו הצליח לסובב את הגלגל מספיק מהר, ולהגיע למהירות סיבוב מינמלית מסוימת, כך שהקרן המוחזרת נתקעה בשן של הגלגל, במקום לעבור שוב דרך אותו חריץ. מתוך הידיעה של מהירות סיבוב הגלגל ואורך המסלול (5 מייל) פיזו הצליח לחשב את מהירות האור שהיא בקרוב $300,000,000$ מטרים לשניה. כיום, במסגרת הסטנדרטיזציה של השיטה המטרית, מהירות האור הוגדרה להיות $299,792,458$ מטרים לשניה.

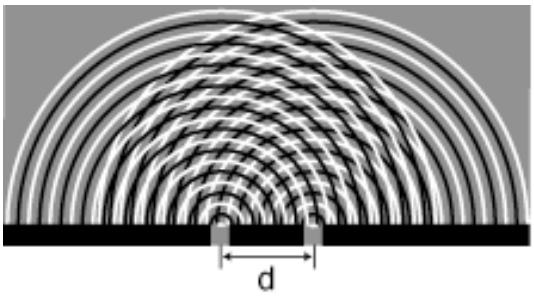


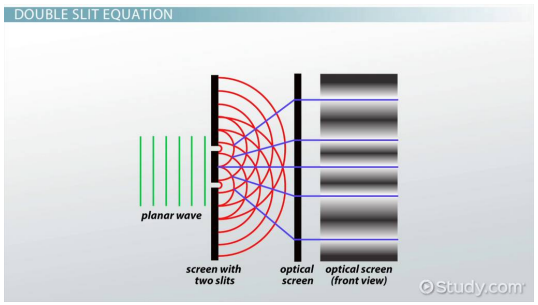
$$[v] = \left[\frac{L}{T}\right] = \frac{meter}{second} \tag{1.8}$$

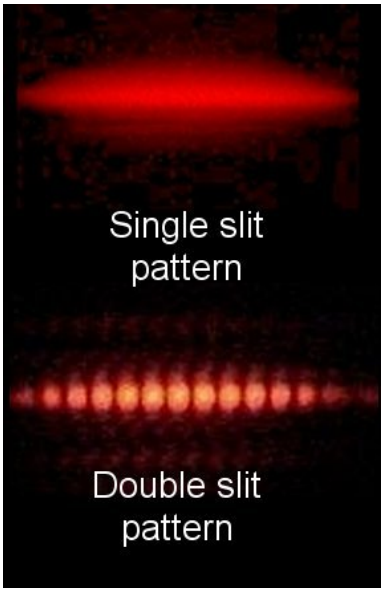
קיימות גרסאות שונות של שיטת פיזו למדידת מהירות. נניח שיש לנו אלומה של אלקטרונים. נשלח את אלומת האלקטרונים לעבר שני דיסקים שמסתובבים על ציר. בהיקף כל דיסק יש חור יחיד, באותה זוית. אם מהירות הסיבוב היא אפס האלומה יכולה לעבור דרך שני החורים. אם מהירות הסיבוב נמוכה הקרן לא עוברת, כי האלקטרונים שעוברים את החור בדיסק הראשון נתקעים בדיסק השני.נגדיל את מהירות הסיבוב עד שנגיע למצב שבו האלומה עוברת. לפי המרחק של הדיסקים וקצב הסיבוב ניתן לקבוע מהי מהירות האלקטרונים.

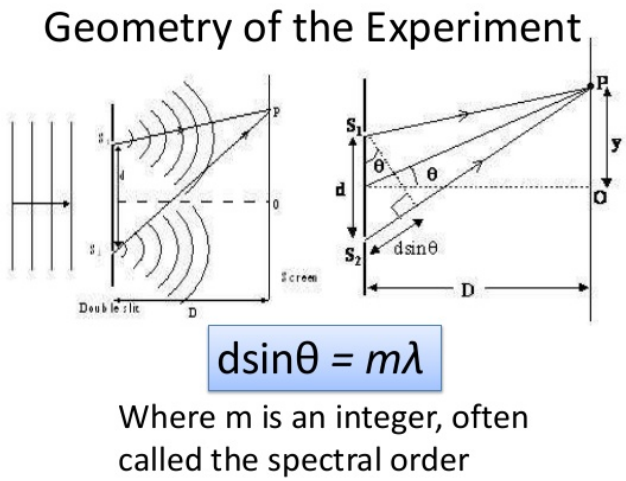
[1.7] Measurement of momentum (wavelength)

בתחילת המאה ה-19 ניסה הפיסיקאי תומאס יאנג להכריע האם האור הוא חלקיק או גל ע"י ניסוי התאבכות: ז"א העברת האור דרך "שני סדקים" או לחילופין אפשר להשתמש "שריג עקיפה". אם האור היה מורכב מחלקיקים היתה נוצרת מריחה אחידה של אור על המרקע. כיוון שהאור הוא תופעה גלית, חיבור תנודות המגיעות ממקורות שונים יוצרת תופעת ההתאבכות:









In the figure above the separation between the stripes of constructive interference on the screen is Δy. The angular separation (in radians) is defined as $\Delta\theta \equiv \Delta y/D$, where D is the distance to the screen. A simple derivation implies that this angular separation is determined by the wavelength λ as follows

$$\Delta\theta = \frac{\lambda}{d} \tag{1.9}$$

where d is the distance between the slits. From this formula we can deduce λ from the measurement of $\Delta\theta$. The momentum (aka wavenumber) is defined as $p = 2\pi/\lambda$.

Derivation [can be skipped].– The intensity of the oscillation that originates from two sources is

$$\text{Intensity} \quad = \quad \left|Ae^{i(\phi_1-\omega t)} + Ae^{i(\phi_2-\omega t)}\right|^2 \quad = \quad 2\left(1+\cos(\phi_2-\phi_1)\right)\ |A|^2 \tag{1.10}$$

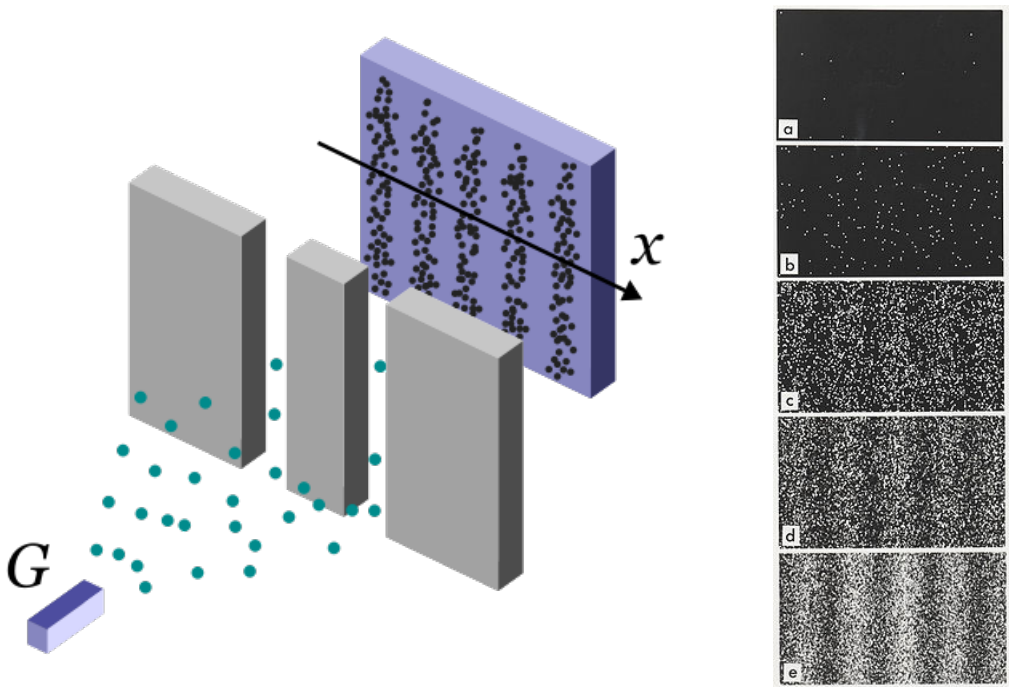
The phase difference between the oscillations that arrive to a given point on the screen is

$$\phi_2 - \phi_1 \quad \approx \quad k \cdot d \cdot \theta \tag{1.11}$$

Hence the angular separation between the interference peaks is $\Delta\theta = \lambda/d$.

═══ [1.8] Momentum measurement of electrons

במהלך המאה ה-20 הומצאה השפופרת הקתודית, מכשיר המאפשר להפיק אלומה של אלקטרונים. בניסוי הדומה לניסוי יאנג אפשר לבדוק את ההתנהגות של אלומת אלקטרונים. אם משגרים את האלקטרונים לעבר המסך דרך חריץ יחיד, התוצאה היא מריחה אחידה של האלקטרונים על גבי המסך. אם יש חריץ נוסף מתקבלת בניגוד לציפיה הקלאסית תבנית מפוספסת המעידה על התאבכות גלית, ואשר מאפשרת לקבוע את "אורך הגל" של האלומה. זה נקרא "אורך גל דה-ברולי".



התוצאה היא מפתיעה. על מנת לעמוד על משמעותה, ולודא שלא מדובר בתופעה קולקטיבית (דוגמת גלי ים), ניתן לשגר לעבר המרקע אלקטרון בודד בכל פעם. מקבלים את אותה התוצאה, בתנאי שצוברים מדגם מספיק גדול של אלקטרונים. שלב נוסף בניסוי הוא הרכבת גלאי על גבי אחד החריצים. בעזרת הגלאי מצליחים לראות דרך איזה חריץ עובר כל אחד מהאלקטרונים. במקרה כזה תופעת ההתאבכות נעלמת, והאלקטרונים מתנהגים באופן "קלאסי".

דואליות גל חלקיק.-- דרך אחת לנסח את המסקנה של הניסוי היא להגיד שלאלקטרון יש תכונה של דואליות: הוא מתנהג כמו חלקיק בזמן שצופים עליו, אך פוגע במרקע ע"פ תופעת ההתאבכות הגלית אם אין מסתכלים עליו לאורך מסלולו. הפגיעות נעשות באופן סטטיסטי - יותר באיזורים המועדים להתאבכות הבונה מאשר באיזורים המועדים להתאבכות ההורסת.

עקרון הסופרפוזיציה.-- האלקטרון, כמו גם הפוטון, וכמו כל חלקיק קוונטי, מתנהגים במהלך תנועתם באופן מוזר - הם כביכול עוברים בשני מקומות בו זמנית כל עוד אין צופים בהם. זה הביא לניסוח הנחת היסוד הבאה של המכניקה הקוונטית: אם שני מצבים הם פיסיקליים אז גם מצב של "סופרפוזיציה" הוא פיסיקלי. במקרה שלפנינו: "להיות בסדק הראשון" זה מצב פיסיקלי, "להיות בסדק

השני" זה מצב פיסיקלי, ולכן גם "להיות בו זמנית בשני סדקים" זה גם מצב פיסיקלי. חלקיק יכול להמצא בעת ובעונה אחת במספר מקומות. במקרה כזה אנו אומרים שאין "ודאות" במיקום של החלקיק.

דטרמיניזם לעומת תאור הסתברותי.-- השאלה המתבקשת היא האם אולי ניתן לתת הסבר "קלאסי" לניסוי, מה שיאפשר באופן עקרוני לחזות איפה יפגע אלקטרון אינדיוידואלי במסך. נניח היפותתית שיש לנו שליטה מלאה באופן שבו אנו משגרים את האלקטרון, ושאנו יודעים עליו כל מה שאפשר לדעת. האם נוכל לקבוע בודאות לאיפה הוא יגיע? ניסוי שני סדקים לא נותן תשובה חד משמעית לשאלה זו. ספקולטיבית אפשר לכאורה להניח שתאוריה עתידית תוכל לקבוע באופן דטרמינסטי את התוצאה של ניסוי מבוקר, כך שלא נצטרך להתפשר על תאור הסתברותי. בניסוחו של אינשטיין "אלוהים לא משחק בקוביות". בהמשך הקורס נראה שקביעתו של אינשטיין היתה מוטעית: הטבע כן משחק בקוביות. העולם שבו אנו חיים אינו דטרמיניסטי. עלינו להסתפק בתאור הסתברותי מסוג זה שנותנת המכניקה הקוונטית.

Einstein's Letter to Max Born (1926).– "You believe in a God who plays dice, and I in complete law and order in a world which objectively exists, and which I in a wildly speculative way, am trying to capture. I firmly believe, but I hope that someone will discover a more realistic way, or rather a more tangible basis than it has been my lot to find. Even the great initial success of the quantum theory does not make me believe in the fundamental dice game, although I am well aware that some of our younger colleagues interpret this as a consequence of senility."

Later paraphrased as "God does not play dice with the world".

[1.9] The dispersion relation

Within the framework of quantum mechanics the Newtonian definition of inertial mass is not appropriate. Instead the mass of a particle is defined in an absolute way, that does not require to fix a reference mass. We shall define mass as a parameter in the "dispersion relation".

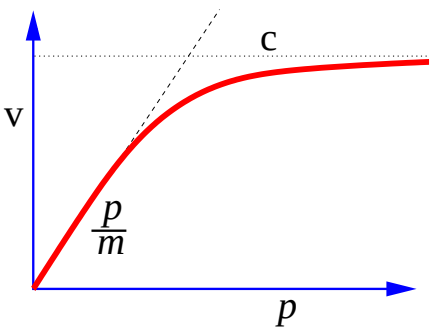
It is possible to prepare a "monochromatic" beam of particles (say electrons) that all have the same velocity. The velocity of the particles can be measured by using a pair of rotating circular plates (discs). It turns out that such beam is further characterized by a wavelength, so-called De-Broglie wavelength. The wavelength of the beam can be measured using a diffraction grating. We define the momentum of the moving particles (aka "wavenumber") as:

$$p \;=\; \frac{d[\text{phase}]}{d[\text{distance}]} \;=\; \frac{2\pi}{\text{wavelength}} \tag{1.12}$$

It is possible to find (say by an experiment) the relation between the velocity of the particle and its momentum. This relation is called the "dispersion relation". Einstein has deduced the following expression

$$v \;=\; \frac{cp}{\sqrt{(\mathsf{m}c^2)^2+(cp)^2}}\,c \tag{1.13}$$

For low (non relativistic) velocities the relation is approximately linear, namely, $v \approx (1/\mathsf{m})p$. Here is a plot of what we expect to observe:



The dispersion relation defines the "mass" parameter. The implied units of mass are

$$[\mathsf{m}] = \frac{T}{L^2} \tag{1.14}$$

If we use arbitrary units for measuring mass, say "kg", then the conversion prescription is:

$$\mathsf{m}[\text{kg}] \quad = \quad \hbar\, \mathsf{m} \left[\frac{\text{second}}{\text{meter}^2}\right], \qquad \hbar = \frac{h}{2\pi} \tag{1.15}$$

where $\hbar$ is known as the Planck constant.

[1.10] Energy [can be skipped]

* ההגדרה של "מסה אינרציאלית" במכניקה קלאסית (באמצעות התנגשויות)
* ההגדרה של "מסה אינרציאלית" במכניקה קוונטית (פרמטר ביחס הדיספרסיה)
* ההגדרה של מסה בתור "אנרגית המנוחה" בתורת היחסות (פצצות אטום...)
* ההגדרה של "מסה גרביטציונית" בתורת הכבידה (באמצעות שקילה / מאזניים / תנועת הכוכבים)

A practical way to define Kinetic Energy in classical mechanics is via the relation

$$v \quad = \quad \frac{\partial E(p)}{\partial p} \tag{1.16}$$

$$E \quad = \quad \sqrt{(\mathsf{m}c^2)^2 + (cp)^2} \quad \approx \quad \mathsf{m}c^2 + \frac{p^2}{2\mathsf{m}} \tag{1.17}$$

Later we shall see that *energy* in quantum mechanics is defined, in analogy with *momentum*, as

$$E \quad = \quad \frac{d[\text{phase}]}{d[\text{time}]} \quad = \quad \frac{2\pi}{\text{TimePeriod}} \tag{1.18}$$

The implied unit are

$$[p] \quad = \quad \left[\frac{1}{L}\right] \quad = \quad \frac{1}{\text{meter}} \tag{1.19}$$

$$[E] \quad = \quad \left[\frac{1}{T}\right] \quad = \quad \frac{1}{\text{second}} \tag{1.20}$$

[1.11] Newton's laws of motion

The "gravitational mass" is defined using a weighting apparatus. Since gravitational theory is not includes in this course, we shall not use that definition. Another possibility is to define "inertial mass". This type of mass is determined by considering the collision of two bodies. We postulates that there is a linear combination of velocities that is conserved in a any collision:

$$p \;=\; \mathsf{m}_1 v_1 + \mathsf{m}_2 v_2 \;=\; \mathsf{m}_1 u_1 + \mathsf{m}_2 u_2 \tag{1.21}$$

where v and u are the velocities before and after a collision. From a collision experiment one can extract the mass ratio of two bodies:

$$\frac{\mathsf{m}_1}{\mathsf{m}_2} = -\frac{u_2 - v_2}{u_1 - v_1} \tag{1.22}$$

In order to give information on the inertial mass of an object, we have to agree on some reference mass, say the "kg", to set the units. (less arbitrary would be to take the mass of the proton as the reference).

The above postulate regrading *momentum conservation* is commonly regarded as Newton's law of motion. Commonly it is decomposed into 3 "laws" that are implied by this postulate. The 1st law concerns a single isolated particle that does not interact with any other particle or field: conservation of momentum implies that it velocity v remain constant. The 2nd law define the notion of *force*. If we have two particles, we define $F_{2\rightsquigarrow 1} = dp_1/dt$, and $F_{1\rightsquigarrow 2} = dp_2/dt$. According to this definition the statement is that force changes that momentum of the particle. To say that the total momentum p remains constant during a collision is like saying that $dp/dt = 0$. The 3rd law is the statement that $dp/dt = F_{1\rightsquigarrow 2} + F_{2\rightsquigarrow 1} = 0$, aka action and reaction.

[1.12] Polarization and Spin

Apart from the degrees of freedom of being in space, the particles also have an inner degree of freedom called "spin". We say that a particle has spin s if its inner degree of freedom is described by a representation of the rotations group of dimension $2s{+}1$. For example, "spin 1/2" can be described by a representation of dimension 2, and "spin 1" can be described by a representation of dimension 3. In particular:

- Electrons have spin $\frac{1}{2}$, hence 180^o difference in polarization ("up" and "down") means orthogonality.
- Photons have spin 1, hence 90^o difference in linear polarizations means orthogonality.

If we position two polarizers one after the other in the angles that were noted above, no particles will pass through. Thus, an abstract mathematical consideration (representations of the rotation group) has very realistic consequences.

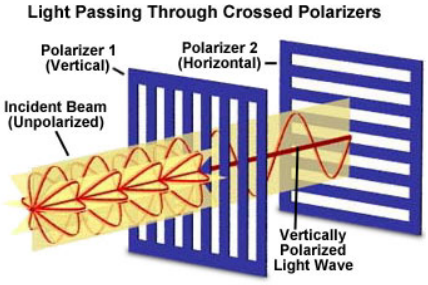
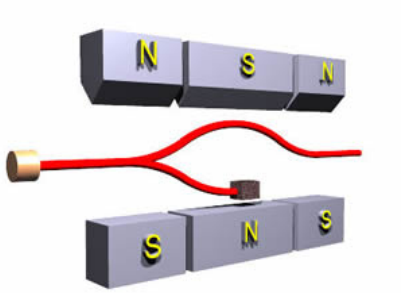


Figure 1



[2] Semiclassical Dynamics

═══ [2.1] The notion of force

Force F is some kind of applied mechanism that affects the dynamics. In stochastic dynamics force is a mechanism to induce drift: we shall explain later that the drift velocity is proportional to the force. But in classical dynamics force is a mechanism to induce change of momentum. It is defined via the relation

$$\frac{d}{dt}\boldsymbol{p} \;=\; F \tag{2.1}$$

For stylistic reasons we use in this section bold letter for the dynamical variables $\boldsymbol{x}$ and $\boldsymbol{p}$. Note that the velocity of the particle is determined by its momentum via a dispersion relation $v = v(\boldsymbol{p})$, as discussed in the Introduction. Related notion is the work W done on a particle by a force F along a trajectory that goes from point A to point B:

$$W_{A\rightsquigarrow B} \;=\; \int_A^B F \cdot dr \tag{2.2}$$

For a so-called conservative force the result is the same for all the trajectories that go from A to B. Then we can define a potential $U(x)$ such that:

$$F(x) \;=\; -\frac{\partial U(x)}{\partial x} \tag{2.3}$$

and we say that $F(x)$ is a field in space, and calculate the work using the formula

$$W_{A\rightsquigarrow B} \;=\; U(x_A) - U(x_B) \tag{2.4}$$

We use here "1D" notations. The generalization to 3D is straightforward.

═══ [2.2] Classical equations of motion

In classical dynamics we have two dynamical variables (x, p) that characterize the state of the particle. It is assumed that the velocity of the particle is determined by a "dispersion relation", namely,

$$\frac{d}{dt}\boldsymbol{x} \;=\; v(\boldsymbol{p}) \;\equiv\; \frac{\partial K(\boldsymbol{p})}{\partial \boldsymbol{p}} \tag{2.5}$$

The function $K(\boldsymbol{p})$ as defined above is called kinetic energy. For non-relativistic particle $v(\boldsymbol{p}) = (1/\mathsf{m})\boldsymbol{p}$ and therefore $K(\boldsymbol{p}) = \boldsymbol{p}^2/(2\mathsf{m})$. The momentum obeys so-called Newton law:

$$\frac{d}{dt}\boldsymbol{p} \;=\; F \;\equiv\; -\frac{\partial U(\boldsymbol{x})}{\partial \boldsymbol{x}} \tag{2.6}$$

We can define Hamiltonian as follows

$$H(\boldsymbol{x}, \boldsymbol{p}) \;=\; K(\boldsymbol{p}) + U(\boldsymbol{x}) \tag{2.7}$$

and derive from it the equations of motion. We define the energy as

$$E = H(\boldsymbol{x}, \boldsymbol{p}) \tag{2.8}$$

It is easily verified that E is a constant of motion ($dE/dt = 0$). The sum of kinetic and potential energies is conserved. Sometimes it is phrased as "the particle changes its kinetic energy due to the work that is done by the force", namely,

$$K(B) - K(A) = W_{A \rightsquigarrow B} \tag{2.9}$$

═══ [2.3] Oscillations

Harmonic oscillator.– This is the term used to describe e.g. particle connected to spring. The force is proportional to the displacement (force constant denote below as α). The outcome is an oscillation with frequency ω that does not depend on the energy (same ω for any amplitude A).

$$x = \text{displacement from equilibrium point} \tag{2.10}$$
$$F = -\alpha x \tag{2.11}$$
$$U(x) = \frac{1}{2}\alpha x^2 \tag{2.12}$$

The equations of motion $\dot{x} = (1/\mathsf{m})p$ and $\dot{p} = F$ lead to Newton's second law $\ddot{x} = (1/\mathsf{m})F$, hence the equation of motion is

$$\ddot{x} = -\frac{\alpha}{\mathsf{m}} x \tag{2.13}$$

whose general solution is

$$x(t) = A\cos(\phi_0 - \omega t) \tag{2.14}$$

The amplitude A and the phase ϕ_0 are determined by the initial conditions at $t = 0$. The frequency is implied by the equation of motion, namely,

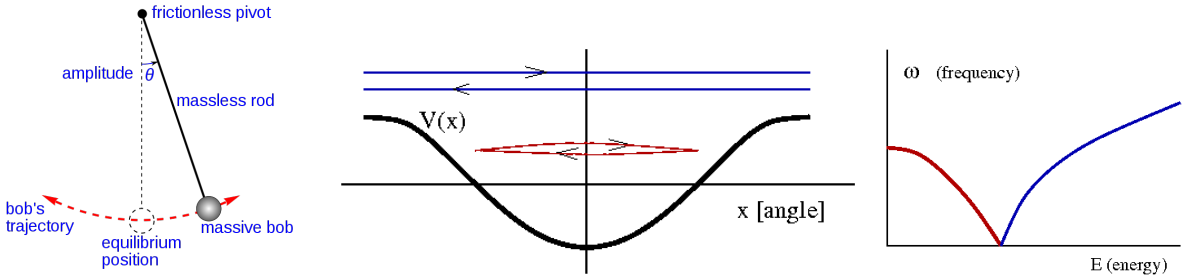
$$\omega(E) = \sqrt{\frac{\alpha}{\mathsf{m}}} \tag{2.15}$$

Square well.– Consider a particle that is free to move in a one-dimensional box of width L. The potential is $U(x) = 0$ inside the box, and very large outside. Given energy E the velocity of the particle is $v = (2E/\mathsf{m})^{1/2}$, the time period of the oscillations is $2L/v$, and therefore

$$\omega(E) = \frac{2\pi}{\text{TimePeriod}} = \frac{\pi}{L} v \propto \sqrt{E} \tag{2.16}$$

Pendulum.– This is an example for unharmonic oscillator. The frequency ω depends on the energy.

$$x = \text{angle from the bottom equilibrium position} \tag{2.17}$$
$$F = -\alpha \sin(x) \tag{2.18}$$
$$U(x) = -\alpha \cos(x) \tag{2.19}$$
$$\omega(E) = \text{[see figure]} \tag{2.20}$$



מרחב הפאזות הוא צילנדרי. נשים לב שעבור תנודות קטנות המטוטלת היא כמו אוסצילטור הרמוני. אבל אם האנרגיה של המטוטלת מספיק גדולה אז היא מבצעת סיבובים שלמים או בכיוון מחוגי השעון או נגד כיוון מחוגי השעון. במילים אחרות: במקרה של מטוטלת יש במרחב הפאזה שלושה אזורים שבכל אחד מהם יש תנועה מסוג אחר.

[2.4] Quantization of the energy

Here we present a naive semi-classical perspective for the quantization of the energy. A prototype example is a particle in a one-dimensional box of length L. The classical energy $E = p^2/(2\mathsf{m})$ can be any positive number $E \in [0, \infty]$. But quantum mechanically p is $2\pi/\lambda$. To have energy E means that the classical is like a standing wave in the box. Clearly this is feasible only if

$$L = (\lambda/2) \times \text{integer} \qquad \text{[semiclassical quantization condition, box]} \tag{2.21}$$

Hence we get

$$|p| = \frac{\pi}{L} \times \text{integer} \tag{2.22}$$

and the energy is quantized

$$E_n = \frac{1}{2\mathsf{m}} \left(\frac{\pi}{L} n\right)^2, \qquad n = 1, 2, 3, ... \tag{2.23}$$

Similar reasoning holds for other bounded systems (e.g. see below – electron in an atom).

Oscillators.– The quantized energies of an harmonic oscillator that has an oscillation frequency Ω is

$$E_\nu = \left(\frac{1}{2} + \nu\right) \Omega, \qquad \nu = 0, 1, 2, ... \tag{2.24}$$

If we take the lower level as reference we get

$$E_\nu \;=\; \nu\Omega, \qquad \nu = 0, 1, 2, ... \tag{2.25}$$

A modes of any harmonic field, e.g. the electromagnetic field, can be in the $\nu = 0, 1, 2, ...$ state. In the modern jargon we say that $\nu = 0$ is the vacuum state, while the $\nu = 1, 2, ...$ are states where the mode occupies ν phonons.

[2.5] Relevance of quantization in practice

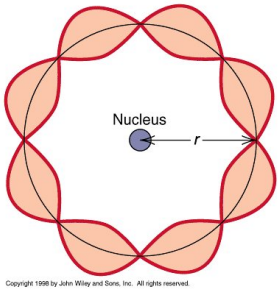
Atoms.– The simplest atom is the Hydrogen atom. The electron has mass m and is attracted to the proton with force $F = -\alpha/r^2$. Assuming the (actually wrong) assumption that the orbital looks circular, we require

$$2\pi r = \lambda \times \text{integer} \qquad \text{[semiclassical quantization condition, circular orbit]} \tag{2.26}$$

which leads to Bohr's quantization condition $\mathsf{m}vr = n$, where $n = 1, 2, 3, ...$ High school physics for circular motion allows, given E, to deduce from Newton's laws that $r = r(E) = -\alpha/(2E)$ and $v = v(E) = \sqrt{-2E/\mathsf{m}}$, see here. Then we can deduce from Bohr's quantization condition that

$$E_n \;=\; -\frac{\mathsf{m}\alpha^2}{2n^2}, \qquad n = 1, 2, 3, ... \tag{2.27}$$

Note that the units are OK (α has dimensions of velocity). This result turns out to be identical to the correct quantum solution of the Schrodinger equation, where the correct shape of the atomic orbitals (s,p,d,f,...) is different.



Energy bands.– Consider a periodic potential, say a sequence of intervals of length a separated by barriers. If the barriers are very high we get energy spectrum like that of a box (with L replaced by a). Each state has L/a degeneracy. If the barriers are lowered, each degenerated level become an energy band. The energy bands are separated by energy gaps.

Thermal occupation.– Consider a system with ordered set of levels E_n, indexed $n = 0, 1, 2, ...$ The state of the system at zero temperature is that of the lowest energy E_0. At finite temperature T the probability to find the system in any of the states is

$$P_n \;\propto\; e^{-(1/T)E_n} \tag{2.28}$$

This reasoning holds irrespective of whether the energy is quantized or not.

[2.6] Concept of Energy in Quantum Mechanics

The definition of energy, say from $v = dE/dp$, implies that the units of energy are 1/sec. In a future lecture we shall see that this is a special case of a more general (exact) formula for the natural frequencies of a system:

$$\omega \;=\; |E_n - E_m| \tag{2.29}$$

Below we list several examples that illustrate this idea.

Square well.– We had argued that the energy of a particle is a box is quantized. It follows that also the frequency ω of the oscillations should be quantized. The naive semiclassical reasoning implies

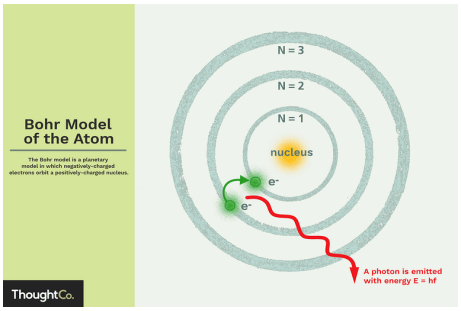
$$\omega[\text{semiclassical}] = \frac{\pi}{L} v_n, \qquad v_n = \frac{1}{\mathsf{m}} \left(\frac{\pi}{L} n\right) \tag{2.30}$$

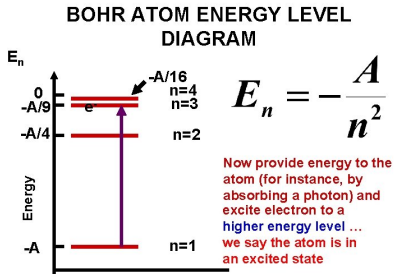
But this is in fact an approximation. The exact way to calculate the natural frequencies is from energy differences. One observed that

$$\omega[\text{quantum}] \;=\; (E_{n+1} - E_n) \;\approx\; \frac{dE_n}{dn} \;=\; \left(\frac{\pi}{L}\right)\frac{dE}{dp} \;=\; \left(\frac{\pi}{L}\right) v \;=\; \omega[\text{semiclassical}] \tag{2.31}$$

Hence we have consistency with the naive expectation.

Spectroscopy.– Illuminating atoms with an irradiation source of frequency Ω, transitions might be induced. The transition from an occupied level E_n that is coupled to a higher level E_m takes place if the resonance condition $\Omega \sim \omega$ is satisfied, where $\omega \equiv (E_m - E_n)$. This is called *absorption*. The inverse process is called *emission*. From quantum mechanical point of view Ω is the frequency of a field mode, meaning that $\Omega = (E_{\nu+1} - E_\nu)$. Hence we regard the resonance condition as an energy conservation law: a photon is absorbed or emitted by the atom.





[2.7] Concept of Resonance

מתנד: לכל מערכת סגורה יש תדירויות אופינית. התדירות מסומנת באות ω, והיחידות הן רדיאן לשניה (אופציונלי יחידות של הרץ - שזה מחזורים לשניה). מבחינים בין מתנד הרמוני (לדוגמה חלקיק קשור לקפיץ) שהתדירות שלו לא תלויה באנרגיה, לבין מתנד לא הרמוני (לדוגמה חלקיק שנע הלוך וחזור בקופסא, מטוטלת) שהתדירות שלו כן תלויה באנרגיה. אופני התנודה (המודים) של שדה (תווך) הם מבחינה מתמטית מתנדים. לדוגמה: תנודות של מיתר, תנודות בלחץ האויר (גלי קול), תנודות של השדה החשמלי (גלי אור: תדירויות שונות = צבעים שונים).

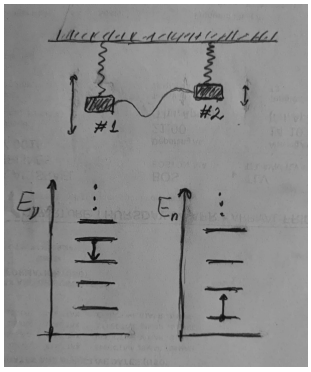
רזוננס - תמונה קלאסית:

כאשר מצמדים שתי מערכות יכולה לעבור ביניהן אנרגיה. לדוגמה: זמרת סופרן שצורחת על כוס זכוכית: אנרגיה עוברת בין התנודות של האויר (גלי קול) לבין התנודות של הכוס (ויברציות של הזכוכית). דוגמה פשוטה יותר: מטוטלות מצומדות. דוגמה עוד יותר פשוטה: גופים תלויים על קפיץ (ראו תרשים). בדוגמה האחרונה הצימוד מיוצג על ידי חוט דק שקושר את שתי המסות. בניסוי אמיתי שאפשר לעשות בבית החוט מיותר כי אנרגיה יכולה לעבור דרך רעידות של הקורה התומכת או דרך האויר.

תנאי הרזוננס הוא התנאי למעבר אנרגיה בין שני הגופים. התנאי הוא שהתדירות של שני הגופים דומה:

$$\omega^{(\#1)} \approx \omega^{(\#2)}$$

אם התדירויות של שני הגופים שונה - האנרגיה לא תצליח לעבור.



רזוננס - תמונה קוונטית:

בתמונה הקוונטית האנרגיה של כל גוף מקוונטטת. הראנו שמרווחי האנרגיה של חלקיק בקופסא שווים בקרוב לתדירות התנועה שנקבעת על ידי מהירות החלקיק במסגרת המכניקה הקלאסית. בעתיד נראה שזה חוק כללי: תדירות התנועה נקבעת לפי מרווחי האנרגיה. אם לדוגמה יש לנו אלקטרון באטום, אז מתקיים

$$\omega^{(atom)} = E_{n+1} - E_n$$

נניח שאנו יוצרים (באמצעות פנס) תנודה של שדה חשמלי. לתנודה הזו יש תדירות כלשהי

$$\omega^{(mode)} = E_{\nu+1} - E_\nu$$

אם למוד התנודה של השדה החשמלי יש אנרגיה E_ν, אנו אומרים שהמוד מאכלס ν פוטונים.

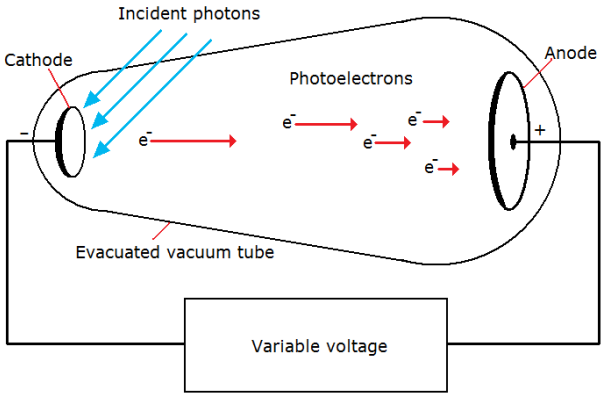
בתמונה הקוונטית אפשר לרשום את תנאי הרזוננס בצורה הבאה:

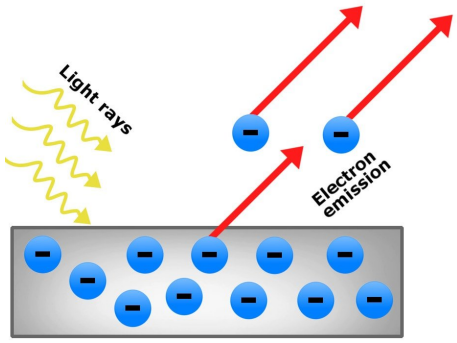
$$E_{\nu+1} - E_\nu \approx E_{n+1} - E_n$$

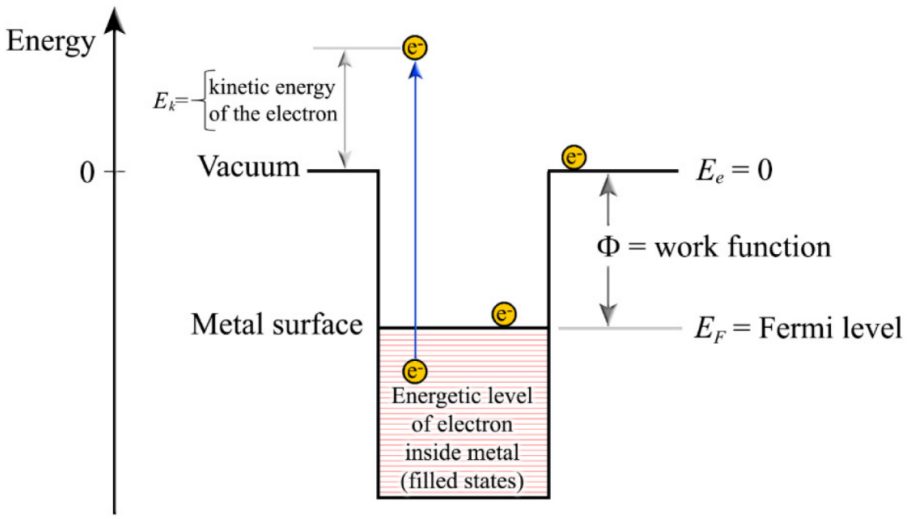
במילים: המוד האלקטרומגנטי מאבד מנה של אנרגיה (פוטון) אשר נבלעת על ידי האטום. האטום מעורר לרמת אנרגיה גבוהה יותר. כמובן שגם העברת אנרגיה בכיוון ההפוך היא אפשרית: האטום המעורר יכול לפלוט חזרה את האנרגיה לשדה האלקטרומגנטי, ז"א לפלוט פוטון. זה הבסיס לספקטרוסקופיה.

[2.8] The photoelectric effect

We can verify experimentally that the quantum perspective of resonance is the correct description. The same considerations as in the story of *absorption* can be used to determine whether a photon of energy ω can ionize an atom. Similarly, the photoelectric effect concerns the emission of electrons from a piece of metal due to irradiation. A threshold frequency is required. Contrary to the classical expectation the intensity does not matter! The kinetic energy of the emitted electrons is $K = \omega - W$, where W is the work function (see figure).







[3] Classical oscillations and waves

[3.1] Oscillators

Back to the harmonic oscillator. Assume that it has an oscillation frequency is ω. The units are rad/sec.

$$T = \frac{2\pi}{\omega} = \text{time period of the oscillation} \tag{3.1}$$

$$f = \frac{1}{T} = \text{frequency in Hz units} \tag{3.2}$$

Denote by q the coordinate of the oscillator. Oscillations with amplitude $\mathcal{A}$ are described by

$$q = \mathcal{A}\cos(\phi_0 - \omega t) \tag{3.3}$$

$$\dot{q} = \mathcal{A}\omega \sin(\phi_0 - \omega t) \tag{3.4}$$

where ϕ_0 is the initial phase. Define a complex coordinate

$$\Psi = q + i\frac{1}{\omega}\dot{q} = \mathcal{A}e^{i\phi} \tag{3.5}$$

The evolution is described by

$$\phi(t) = \phi_0 - \omega t \tag{3.6}$$

Accordingly

$$\psi(t) = Ae^{-i\omega t} \tag{3.7}$$

We have absorbed the initial phase ϕ_0 into A, namely

$$A \equiv \mathcal{A}e^{i\phi_0} \tag{3.8}$$

The complex amplitude A is determined by the initial conditions: the real part by the initial position $q(0)$, and the imaginary part by the initial velocity $\dot{q}(0)$. The motion is a clockwise trajectory along a circle in the complex plane, and we have the relation

$$q(t) = \Re\left[Ae^{-i\omega t}\right] \tag{3.9}$$

[3.2] Coupled oscillators

The equation of motion for N coupled oscillators can be written in a matrix form:

$$\ddot{\boldsymbol{q}}_i = -\sum_{j=1}^{N} \boldsymbol{W}_{ij}\boldsymbol{q}_j \tag{3.10}$$

We look for the *modes* of the system. By definition, each mode of motion is described by a single frequency ω. We substitute the desired form of solution:

$$\boldsymbol{q}_j = \Re\left[\boldsymbol{A}_j e^{-i\omega t}\right] \tag{3.11}$$

where the A_j are complex amplitudes. We get the equation

$$\boldsymbol{W}\boldsymbol{A} = \omega^2 \boldsymbol{A} \tag{3.12}$$

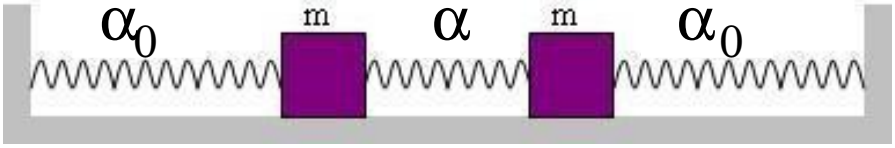
From that we find N independent solutions. The general solution $\boldsymbol{q}(t) = \Re[\Psi(t)]$ is a superposition of the modes:

$$\Psi(t) = \sum_{n=1}^{N} c_n \boldsymbol{A}^{(n)} e^{-i\omega_n t} \tag{3.13}$$

The coefficients c_n are determined by the initial conditions.

═══ [3.3] Modes of masses coupled by springs

We first consider two masses coupled by spring as in the following figure:



Here we have $N = 2$ masses. The $\boldsymbol{W}$ matrix is

$$\boldsymbol{W} = \frac{1}{\mathsf{m}}\begin{pmatrix} \alpha_{\text{eff}} & -\alpha \\ -\alpha & \alpha_{\text{eff}} \end{pmatrix} \tag{3.14}$$

where $\alpha_{\text{eff}} = \alpha_0 + \alpha$. The frequencies of the two modes are

$$\omega_\pm = \sqrt{\frac{\alpha_{\text{eff}} \mp \alpha}{\mathsf{m}}} \tag{3.15}$$

The modes are

$$\boldsymbol{A}^{(+)} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{3.16}$$

$$\boldsymbol{A}^{(-)} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{3.17}$$

The normalization of the eigenvectors is a matter of convention. It is more convenient to use Dirac notations. Those notation generalizes the usual vector notations. Namely, $\vec{A} = \vec{B} + \vec{C}$ is written as $|A\rangle = |B\rangle + |C\rangle$. Thus we write

$$|+\rangle = \frac{1}{\sqrt{2}}\left[|1\rangle + |2\rangle\right] \tag{3.18}$$

$$|-\rangle = \frac{1}{\sqrt{2}}\left[|1\rangle - |2\rangle\right] \tag{3.19}$$

The $|+\rangle$ mode has a frequency $\omega_+ = \sqrt{\alpha_0/\mathsf{m}} \equiv \omega_0$ that does not depend on α, because the masses are oscillating with the same phase, such that the α spring is not involved. The general solution is any linear combination of the two modes:

$$|\psi(t)\rangle \;=\; \sum_{\pm} c_n e^{-i\omega_n t}\,|n\rangle \;=\; c_+ e^{-i\omega_+ t}\begin{pmatrix}1\\1\end{pmatrix} + c_- e^{-i\omega_- t}\begin{pmatrix}1\\-1\end{pmatrix} \tag{3.20}$$

Let us generalize the solution for the case where we have large N. The equilibrium location of mass j mod(N) is $x = ja$, where a is the so called lattice constant. The $\boldsymbol{W}$ matrix takes the form

$$\boldsymbol{W} \;=\; \begin{pmatrix} \alpha_{\text{eff}} & -\alpha & 0 & ... & -\alpha \\ -\alpha & \alpha_{\text{eff}} & -\alpha & 0 & ... \\ 0 & -\alpha & \alpha_{\text{eff}} & -\alpha & ... \\ ... & ... & ... & ... & -\alpha \\ -\alpha & 0 & ... & -\alpha & \alpha_{\text{eff}} \end{pmatrix} \;=\; \alpha_{\text{eff}}\mathbf{1} - \alpha\boldsymbol{D} - \alpha\boldsymbol{D}^{-1} \tag{3.21}$$

For mathematical simplicity we assume periodic boundary condition (a closed chain). Note that $\alpha_{\text{eff}} = \alpha_0 + 2\alpha$ instead of $\alpha_{\text{eff}} = \alpha_0 + \alpha$ because each mass is connected to two α springs. We have expressed $\boldsymbol{W}$ in terms of the displacement operator $\boldsymbol{D}$. By definition $\boldsymbol{D}\,|x\rangle = |x{+}a\rangle$. We see that $\boldsymbol{W}$ will be diagonal in the same basis where $\boldsymbol{D}$ is diagonal. Therefore to find the eigenvector is an easy task. These are plane waves

$$|k\rangle \;=\; \frac{1}{\sqrt{N}}\sum_x e^{ikx}\,|x\rangle \tag{3.22}$$

such that $\boldsymbol{D}\,|k\rangle = e^{-ika}\,|k\rangle$. The corresponding frequencies are

$$\omega_k \;=\; \sqrt{\frac{1}{\mathsf{m}}\Big[\alpha_{\text{eff}} - 2\alpha\cos(ka)\Big]} \;=\; \sqrt{\omega_0^2 + \frac{\alpha}{\mathsf{m}}\cdot 2[1-\cos(ka)]} \tag{3.23}$$

where $\omega_0 = \sqrt{\alpha_0/\mathsf{m}}$. Of particular interest is the case where $\alpha_0 = 0$. It means that the masses are inter-connected with the α springs, but not to the ground. Then, for small k one obtains $\omega_k \approx ck$, where $c = (\alpha/\mathsf{m})^{1/2}a$.

Polarization.– Above we had masses that are distinguished by their locations. The notation $|x\rangle$ means that we have a displaced mass in location that is labelled by x. In general the mass can be displaced in any direction. So we need an additional label s to indicate the so-called polarization. Thus we use the notation $|x,s\rangle$ in order to indicate displacement at location x in direction s.

═══ [3.4] Digression - diagonalization by inspection

We have encountered above the first Pauli matrix $S \equiv \sigma_1$ and the displacement matrix D. Let us write these matrices and see how they can be diagonalized by inspection:

$$S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \boldsymbol{W} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \tag{3.24}$$

Note that S performs swap. We can label the two basis vectors as $x = 0, 1$. They can represent two sites or two states of a qubit. The eigen-vectors and the eigen-values are determined by the equation $S\psi = s\psi$. By inspection they are

$$|s{=}1\rangle = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad |s{=}-1\rangle = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{3.25}$$

A more compact notation for the same eigen-vectors is

$$|k\rangle = \begin{pmatrix} 1 \\ e^{ik} \end{pmatrix} = \sum_{x=0,1} e^{ikx} |x\rangle, \qquad k = 0, \pi \tag{3.26}$$

We turn to diagonalize by inspection the displacement matrix. Here we label the basis vectors as $x = 0, 1, 2, \cdots, N{-}1$. Note that x is defined $\mod (N)$. In the example above $N = 5$. The transformation $\tilde{\psi} = D\psi$ in matrix notation is written as follows:

$$\begin{pmatrix} \tilde{\psi}_0 \\ \tilde{\psi}_1 \\ \tilde{\psi}_2 \\ \tilde{\psi}_3 \\ \tilde{\psi}_4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \psi_0 \\ \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = \begin{pmatrix} \psi_4 \\ \psi_0 \\ \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix} \tag{3.27}$$

Or it can be written as $\tilde{\psi}_x = \psi_{x-1}$. By inspection, in order to obtain an eigenvector, the phase-difference between subsequent sites should be the same, namely,

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ e^{ik} \\ e^{i2k} \\ e^{i3k} \\ e^{i4k} \end{pmatrix} = \begin{pmatrix} e^{i4k} \\ 1 \\ e^{ik} \\ e^{i2k} \\ e^{i3k} \end{pmatrix} = e^{-ik} \begin{pmatrix} e^{i5k} \\ e^{ik} \\ e^{i2k} \\ e^{i3k} \\ e^{i4k} \end{pmatrix} = e^{-ik} \begin{pmatrix} 1 \\ e^{ik} \\ e^{i2k} \\ e^{i3k} \\ e^{i4k} \end{pmatrix} \tag{3.28}$$

The last equality requires $5k = 0 \mod (2\pi)$, that can be written as $k = (2\pi/N) \times \text{integer}$, or as $\lambda \times \text{integer} = N$, where $\lambda = 2\pi/k$ is the wavelength. These are the permitted wavelengths that satisfy the periodicity of the system. In compact notation we write (normalization does not matter):

$$|k\rangle = \sum_{x=0}^{N-1} e^{ikx} |x\rangle, \qquad k = \frac{2\pi}{N} \times \text{integer} \tag{3.29}$$

[3.5] The term "wavefuntion"

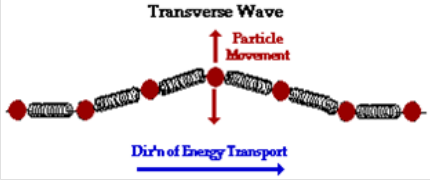
We have defined a complex coordinate Ψ in order to describe the state of an oscillator. If we have several oscillators we use the notation Ψ_j. If we have an oscillator in any location x we use the notation $\Psi(x)$. Namely, $\Psi(x)$ is the "displacement" of the oscillator that is located at x. We say that $\Psi(x)$ describes a field.

It is important to realize that from a mathematical point of view there are different terms that can be used for $\Psi(x)$. In Linear Algebra is is regarded as a *vector* with representation index x. In Calculus in is regarded as a *function* that associates a complex value Ψ to each point x is space. From Mathematical Physics point of view it describes the *state* of a physical field.

Fields that are described by simple Hamiltonians of the type that we have discussed change in time. In particular there are field configurations that oscillate in a well defined frequency ω. These are the *modes* of the field. In general $\Psi(x)$ can be any superposition of those modes.

[3.6] Waves

Consider a chain of identical harmonic oscillators. Each is located at a different position.



Let us assume that they oscillate with frequency ω. Each oscillator may have a different initial phase. Accordingly we write $\phi(x,t) = \phi_0(x) - \omega t$. If the oscillations have a fixed phase-lag between adjacent locations write

$$\phi(x,t) = kx - \omega t \tag{3.30}$$

where k is called wavenumber. Accordingly we get a wave that is described by the wave-function

$$\Psi(x,t) = Ce^{i(kx-\omega t)} \tag{3.31}$$

where C is the amplitude. In order to avoid accidents with notations, we reserve the letter A for indicating the normalized eigen-vectors $\boldsymbol{A}^{(k)}$. Here $A_x^{(k)} = N^{-1/2}e^{ikx}$ with associated eigen-frequencies $\omega = \omega_k$. We define wavelength as the distance where oscillation have the same phase

$$\lambda = \frac{2\pi}{k} = \text{wavelength} \tag{3.32}$$

The relation between the frequency and the wavenumber is called dispersion relation. Electromagnetic waves and sound waves have a linear dispersion relations

$$\omega_k = c|k| \tag{3.33}$$

We shall argue below that c is, in some sense, the propagation velocity of the wave.

[3.7] Wave-packets

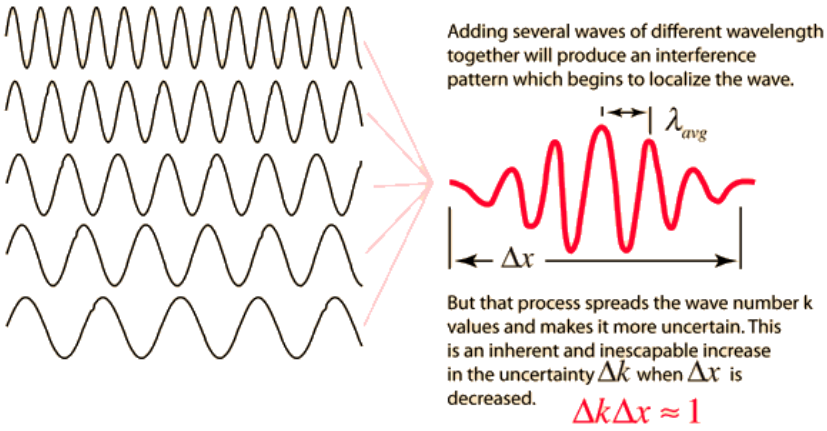
Striking a string, a wave-packet is produced. The wavepacket has a spatial width Δx, and involves a superposition of plane waves, with dispersion Δk. Namely,

$$\Psi(x,t) \;=\; \sum_k C_k e^{i(kx-\omega_k t)} \tag{3.34}$$

The superposition coefficients C_k are determined (via Fourier analysis) by the initial condition at $t=0$. Namely, by the initial shape $f(x)$ of the pulse as follows:

$$\Psi(x,0) \;\equiv\; f(x) \;\equiv\; \sum_k C_k e^{ikx} \tag{3.35}$$

After some time the wavepacket spreads over the whole string (if the string is of finite length). It is also possible to produce a standing wave via superposition of k and $-k$ terms that have the same frequency ω_k.



For a wave with linear dispersion relation we can write

$$\Psi(x,t) \;=\; f(x-ct) \tag{3.36}$$

More generally we can linearize any dispersion relation around the central frequency and deduce

$$v \;=\; \frac{d\omega}{dk} \qquad \text{[calculated for the average wavenumber]} \tag{3.37}$$

An example for a non linear dispersion relation that describes balls that are connected to the ground by springs, and also inter-connected by flexible bonds:

$$\omega_k \;=\; \sqrt{\omega_0^2 \;+\; (ck)^2} \tag{3.38}$$

It turns out that also a beam of electrons has formally the same dispersion relation, where k is called momentum (denoted by p), and ω_0 is called rest-energy (denoted by mc^2). What is the meaning of "Ψ" in the latter context will be discussed in a later lecture.

[4] Electricity

═══ [4.1] The electric field

It turns out that particles can influence each other via so-called electrostatic interaction that is described by Coulomb law. The force is written as $F = Qq/r^2$, where r is the distance, while Q and q denote the electric *charge* of the particles. We prefer to regard this interaction in asymmetrical way: one charge (Q) is creating the field, and the other (q) is influenced by the field. Accordingly we write the electric field of a point charge Q as $\mathcal{E} = Q/r^2$, and the force on the the other particle as

$$F = q\mathcal{E} \tag{4.1}$$

The electric field $\mathcal{E}$ for different charge configuration can be deduced from Coulomb law, but it is more convenient to say that it is determined by Gauss law:

$$\text{Flux of Electric Field} \equiv \oiint \mathcal{E} \cdot dS = 4\pi\, Q \tag{4.2}$$

For a point charge Q we get

$$\mathcal{E} = \frac{Q}{r^2} \tag{4.3}$$

while for a large plate of area A we get

$$\mathcal{E} = 2\pi\frac{Q}{A} \tag{4.4}$$

perpendicular to the plate (on each side).

We can define an electric potential $V(x)$, such that

$$\mathcal{E}(x) = -\frac{d}{dx}V(x) \tag{4.5}$$

For a point charge we get

$$V(r) = \frac{Q}{r} \tag{4.6}$$

For two parallel plates, with distance d separation, and opposite charges $\pm Q$, we get potential difference

$$V = 4\pi\frac{Qd}{A} \tag{4.7}$$

The dynamics of single particle of charge q is determined by Hamiltonian with the potential energy

$$U(x) = qV(x) \tag{4.8}$$

[4.2] The RC circuit

Consider a metallic sphere of radius r_0 with charge Q. Its potential relative to the "ground" at infinity is $V = Q/r_0$. We write it as $V = Q/C$ where the capacitance is $C = r_0$. We can define C for other geometries. For a plate capacitor V is the potential difference between the plates (optionally one can regard one plate as "grounded"), and we get

$$C = \frac{A}{4\pi d} \tag{4.9}$$

where A is a the area and d is the distance between the plates. Ideally, the work to charge a capacitor is

$$W = \frac{Q^2}{2C} \tag{4.10}$$

This formula can be proved in two ways. One way is to integrate $dW = V(q)dq$ from $q = 0$ to $q = Q$, where $V(q) = q/C$. The other way, specifically for a plate capacitor, is to calculate the work $W = Fd$ that has to be done in order to separate the two plates (starting from zero distance), where the force is $F = 2\pi Q^2/A$.

We can discharge the capacitor. For this purpose we connect a conductor of resistance R between the sphere and the ground (or between the two plates). The current is given by Ohm law

$$I = \frac{1}{R}V \tag{4.11}$$

The equation that describes the circuit is

$$\frac{d}{dt}Q = -I = -\frac{V}{R} = -\frac{1}{RC}Q \tag{4.12}$$

The solution is

$$Q(t) = Q(0)e^{-t/\tau} \tag{4.13}$$

where $\tau = RC$ is the "time constant" of the decay. Note that the work that is done on charge $dQ = Idt$, that is transferred through the resistor, is

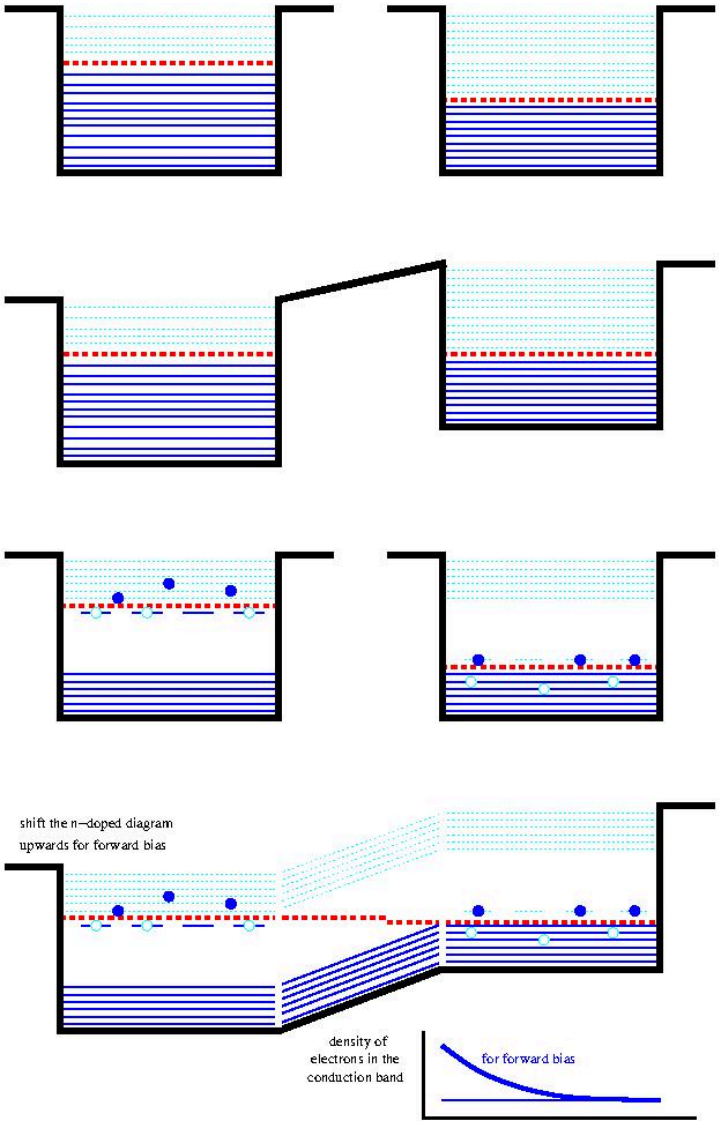
$$dW = VdQ = RI^2dt \tag{4.14}$$

This is known as Joule law. The work increases the kinetic energy of the electrons, but at the same time this kinetic energy is dissipated as heat, or possibly emitted as electromagnetic energy (light).

[4.3] Metals and Semiconductors

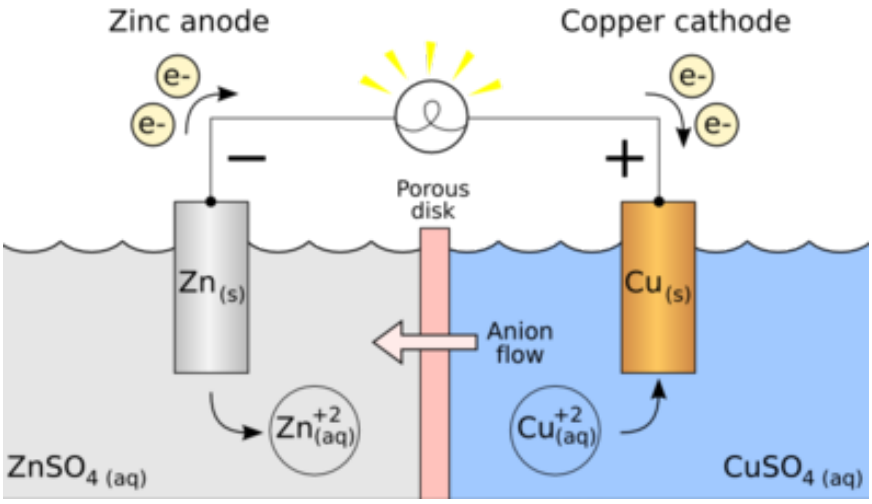
מבנה הפסים הטיפוסי של מוליך למחצה כולל פס ערכיות שיכול לאכלס חורים, פער אנרגטי (gap) , ופס ערכיות שיכול לאכלס אלקטרונים. במוליך למחצה "אינטרינסי" מספר האלקטרונים שווה למספר החורים, והכמות של נושאי המטען נקבעת על ידי הטמפרטורה. אם מוסיפים זיהומים $(donors, acceptors)$, אפשר להגדיל את כמות החורים או את כמות האלקטרונים, ובהתאם מדברים על מוליך למחצה מטיפוס p או על מוליך למחצה מטיפוס n . אם יוצרים על-פני מוליך למחצה מטיפוס p שער $(gate)$, אז פוטנציאל חיובי יגרום להתרוקנות $(depletion)$ מנושאי מטען כיוון שהחורים "יידחו" מהשער. אבל אם הפוטנציאל מאוד חזק תיווצר שכבת היפוך $(inversion\ layer)$ שכוללת נושאי מטען שליליים.

Figure: **(a)** Two different metals with the same electrostatic potential. Electrons flow from the metal that has the larger chemical potential $(\mu_L > \mu_R)$. **(b)** Two different metals at equilibrium. Same chemical potential. No flow of current. In order to get current (not displayed) we have to induce a potential difference $((\mu_L - \mu_R) = eV)$. **(c)** Non coupled n-doped and p-doped semiconductors. **(d)** The np junction at equilibrium $(\mu_L = \mu_R)$. At zero temperature electrons form the n-side donors are depleted into the p-side acceptors. A depletion region is formed, and a constant potential is established. In order to get current we have to induce a potential difference. The plot illustrated the electron density at the p-side. Finite temperature is assumed (otherwise the carrier density is zero).



[4.4] The design of a Battery

The ions in the solution equalize the electrostatic potential (metal Z_n^+ cations dissociate from one electrode; Cu^+ cations join the other electrode; and SO_4^{--} flow to counter ballance the inter-cell potential difference). Hence contact potential difference is not established, and the flow of electrons persists.



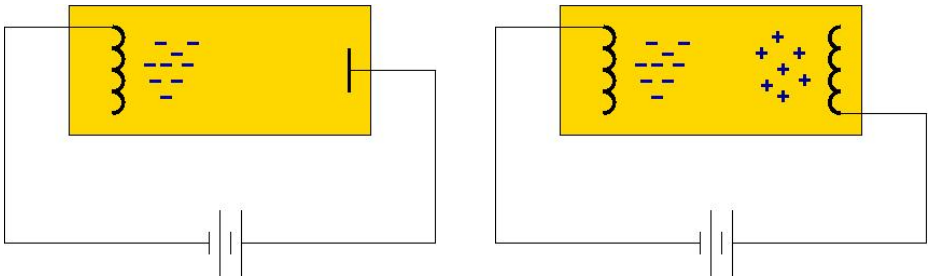
[4.5] The Diode

Ohm law $I = GV$, where the conductance is $G = 1/R$, is a simple example for "stochastic dynamics" that we discuss in the next section: the drift velocity of the electrons is proportional to the applied field. A more sophisticated example for stochastic dynamics is provided by the Shockley diode equation

$$I = I_S \left[e^{|q|V/T} - 1 \right] \tag{4.15}$$

where T is the temperature and $q = -|q|$ is the charge of the electron. It is a matter of convention to define $V > 0$ as the forward bias. Note that for very small V we have approximately Ohmic behaviour with $G = (|q|/T)I_S$, but for reversed bias ($V < 0$) the current saturates at a very small value I_S, unlike forward bias for which the current can be very large.

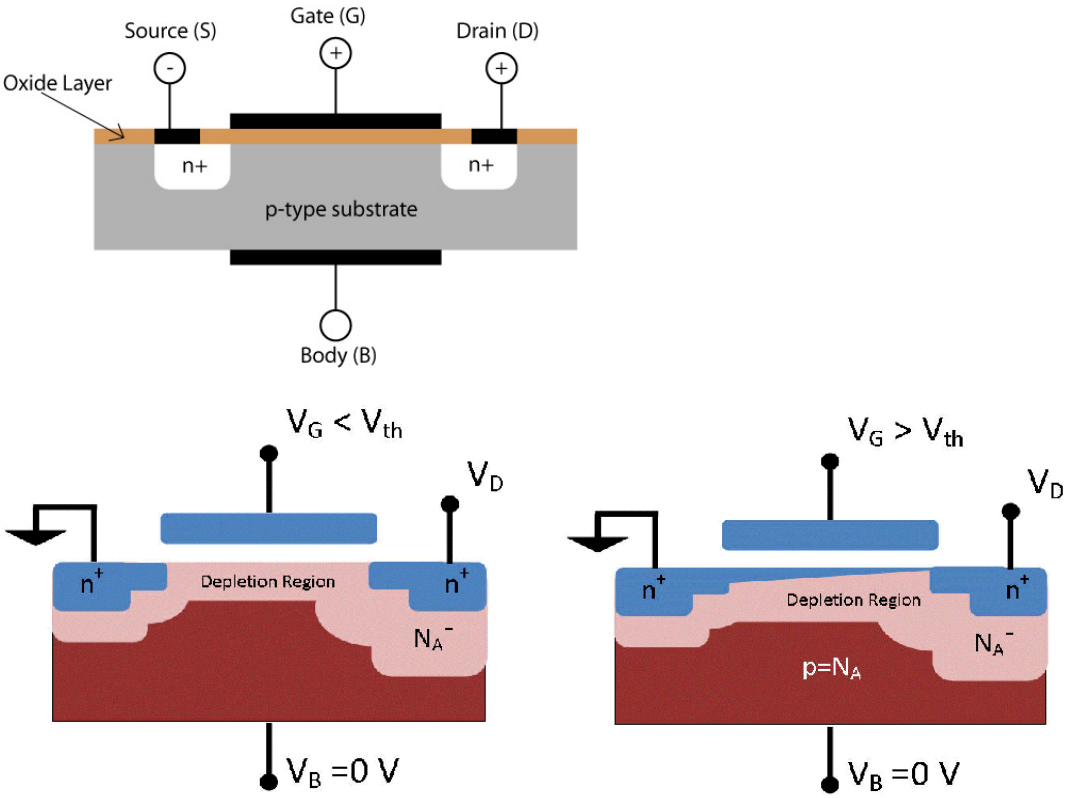
The derivation of the Diode equation goes as follows: (1) The thermionic emission of a metal goes like $I \propto e^{-W/T}$ where W is the work function. (2) If we attach two metals with different W-s we get a current, until it is counter-balanced by a contact potential. (3) Similarly, if we attach n-doped semiconducator with p-doped semiconducator a depletion layer is formed, which implies a potential step W that equals the band gap. (4) If we add forward bias, the potential at the p becomes lower, namely, it equals $W+U$ with negative $U = qV$. Accordingly the electrons at the n-side face a lower potential step, and their density at the p-side of the depletion layer is enhanced by factor $e^{|q|V/T}$. (5) Looking at the p-side the current of electrons is proportional to $(e^{|q|V/T} - 1)$. This is the minority current at the p side and it should be equal to the majority current of the same electrons at the n side. (6) The same considerations apply if we look on the hole current. The total current is the sum of electron and hole currents across the junction. Note: this derivation assumes "narrow junction" meaning that there is no enough time for recombination within the dpeletion region.



להלן הסבר אינטואיטיבי לפעולת הדיודה. על מנת להבין את עקרון הפעולה של הדיודה נתיחס תחילה לגרסא השפופרתית שלה שכוללת קתודה ואנודה כמתואר בשרטוט. בממתח חיובי האלקטרונים שנפלטים מהקטודה מואצים לעבר האנודה ונוצר זרם. בממתח הפוך כמעט אין זרם. התקן כזה נקרה "יוניפולרי" כיוון שהוא מבוסס על סוג אחד של נושאי מטען (אלקטרונים). אם אפשר היה ליצור אנודה שפולטת נושאי מטען חיוביים (פוזיטרונים), אז היה מתקבל התקן "ביפולרי". אם שמים ממתח חיובי אז במרכז ההתקן יהיה אזור "רקומבינציה", כך שהזרם מהקתודה עד לאיזור הזה הוא בעיקר עם אלקטרונים, בעוד שהזרם באיזור של האנודה הוא בעיקר עם פוזיטרונים. התקן ביפולרי ממומש בפועל על ידי הצמדת מוליך למחצה מסוג n למוליך למחצה מסוג p. זה נקרא $pn\ junction$. אנרגית הרקומבינציה של האלקטרונים והחורים בצומת יכולה להשתחרר בצורת אור, ואז מדברים על דיודה פולטת-אור LED.

[4.6] The transistor

אפשר לתאר טרנזיסטור כרכיב אלקטרוני שאפשר לשלוט על ההתנגדות שלו באמצעות שער $(gate)$. בדומה לשליטה על הזרם בצינור מים באמצעות ברז, זה מאפשר ליצור הן מעגלי מיתוג דיגיטליים, והן מעגלי הגברה אנלוגיים (במקרה האחרון שינויים קטנים במתח השער שולטים על זרם גדול). ישנן מספר דרכים לבנות טרנזיסטור, אך נבחר להרחיב על המבנה של טרנזיסטור מסוג $MOSFET$, כמתואר בשרטוט. אם נתעלם מה- $gate$ הרי שיש לנו שתי דיודות מחוברות גב אל גב ולכן לא יכול לזרום זרם בין ה- $source$ לבין ה- $drain$. על מנת לקבל זרם יש לשים את ה- $gate$ בממתח חיובי מספיק גדול כך שתיווצר שכבת היפוך.



[4.7] Electronics

דיודות וטרניזסטונים הם אבני הבניין של האלקטרוניקה - תחום שהתפתח במהלך המאה ה-20. דיודה יכולה לשמש כדי "ליישר" זרם חילופין ולהפוך אותו לזרם ישר. יישור המתח יכול להיות עבור אספקת מתח למכשירים חשמליים, וגם יכול להיות כדי להפריד את הסיגנל שמשודר למקלט רדיו על גבי "גל נושא" בעל תדירות גבוהה. טרניזסטורים יכולים לשמש במעגלים דיגיטליים לצורך בניית שערים לוגיים. בהקשר אחר - טרנזיסטורים יכולים לשמש לצורך הגברה של אות אנלוגי: אות חלש שמוזן לשער של הטרנזיסטור שולט על זרם חזק בין ה- $source$ לבין ה- $drain$. מוליכים למחצה בקונפיגורציות שונות (מוליך/דיודה) יכולים לשמש כגלאים לאור עבור מצלמות וכיו"ב.

[4.8] Superconductors

פוטונים הם בוזונים - אפשר לאכלס הרבה פוטונים באותו מוד (אותו מצב תנע עם אותו קיטוב) - אלומה של לייזר. אפשר לעשות משהו דומה עם אטומים קרים (קונדסציית בוזה-אינשטיין). אלקטרונים הם פרמיונים - חוק האיסור של פאולי. אי אפשר לשים שני אלקטרונים באותו מצב חד-חלקיקי (אותו מצב תנע עם אותו מצב ספין). זוגות קופר נוצרים בגלל משיכה אפקטיבית שמתווכת (בסופר מוליך קונבנציונלי) על ידי השריג. הזוגות עושים קונדסציה במצב תנע מסוים (נניח תנע אפס). בדומה למוליך למחצה יש gap לעירורים. בטמפרטורה נמוכה ארוע של שבירת זוג היא בעלת סבירות אפסית. הסבר של תופעת הסופר מוליכות (תהליך של העברת החלקיקים מקונדסציה במומנטום אחד לקונדסציה במומנטום אחר - סבירות אפסית). בצומת ג'וזפסון מצבי הבסיס "n" הם מצבים קוונטיים מוגדרים היטב. בניגוד למצבי "n" של קבל לוחות רגיל.

[4.9] Josephson Junctions

המימוש הפשוט ביותר של קיוביט מבחינה קונצפטואלית זה באמצעות "ספין" של חלקיק. במקרה כזה מצבי הקיוביט הם מצבי קיטוב - על זה נדבר בהמשך הקורס. אפשרות אחרת, זהה מבחינה פורמאלית, היא לממש מערכת שיש לה שני מצבי בסיס. לכאורה "חלקיק במערכת שני אתרים" יכול להתאים למשימה. אבל האפשרות הזו מאתגרת מבחינת פבריקציה. האתר צריך להיות מאוד קטן $(quantum\ dot)$ על מנת שיהיה אפשר להזניח (בטמפרטורות עבודה מעשיות) עירורים לרמות אנרגיה גבוהות. כיום הטכנולוגיה המובילה היא מימוש של קיוביט באמצעות צומת ג'וזפסון. צומת כזו היא מעין קבל שהלוחות שלו הם סופר-מוליכים. במימוש המקובל $(charge\ qubit)$ מצבי הבסיס "n" של של הצומת נבדלים במספר זוגות הקופר שעוברים מלוח אחד ללוח שני. כדי שההתקן יעבוד הטמפרטורה צריכה להיות יותר קטנה מהאנרגיה שדרושה לפרק זוג-קופר. זה תנאי שקל יחסית לממש באופן פרקטי.

[5] Stochastic Dynamics

[5.1] The emergence of Stochastic dynamics

The dynamics of a closed isolated particle is described by the quantum Hamiltonian or its classical approximation. If the particle is not isolated from the environment the dynamics becomes stochastic. In classical language we say that the interaction with the thermal environment diminishes the average velocity of the particle. In quantum language we say that the interaction with the thermal environment diminishes the coherences. Consequently we can describe the dynamics in a reduced phase-space where only the x coordinate matters.

Consider, without loss of generality, two sites n and m that have binding energies E_n and E_m respectively. In the absence of field a thermal environment is likely to induce transitions with some rate $w_{nm} = w_{mn}$. In the presence of field the transitions will be biased. It turns out that

$$\frac{w_{mn}}{w_{nm}} \;=\; \exp\left[\frac{E_n - E_m}{T}\right] \tag{5.1}$$

where T is called temperature (by this definition it has units of energy).

At thermal equilibrium the probabilities P_n and P_m satisfies the so-called detailed-ballance relations

$$w_{mn}P_n \;=\; w_{nm}P_m \tag{5.2}$$

and we deduce that

$$P_n \;\propto\; e^{-(1/T)E_n} \tag{5.3}$$

More generally, if we have many *many* energy states E_n, the same reasoning applies. This is called *canonical* thermal state.

[5.2] Drift and diffusion

Consider transitions between sites of a chain, that are biased by a field F. Given that the spacing between sites is a, and the temperature is T, we deduce that the ratio between forward and backward transitions is

$$\frac{w^+}{w_-} \;=\; \exp\left[\frac{Fa}{T}\right] \tag{5.4}$$

If F is weak, or if a is small (aka "continuum limit") we get that $w^+ \sim w^- \equiv w$, while $w^+ - w^- \approx [(F\,a)/T]w$. Below we further analyse the stochastic dynamics of a particle along a chain. From the analysis we conclude that in the presence of a weak uniform field F the drift velocity is

$$\frac{d}{dt}\langle x\rangle \;=\; \mu\, F \tag{5.5}$$

where $\mu = wa^2/T$ is the so-called mobility, expressed it in terms of the zero field transition rate w, and the lattice constant a. In contrast with isolated system, here the force induces drift rather than acceleration. We shall also

show that the mobility is related to the diffusion coefficient via Einstein's relation:

$$\mu \;=\; \frac{1}{T}D \tag{5.6}$$

The diffusion coefficnet $D = wa^2$ determines the spreading via $\mathrm{Var}(x) = 2Dt$.

======= [5.3] Two level/site system

A two site system is the simplest setup for illustration of equilbration process. The sites are labeled as #1 and #2. Let us assume that the rate w^+ of transitions from site #1 and #2 is larger than the rate w^- of transitions from site #2 and #1. Later we use the notations $\gamma = w^+ + w^-$ and $u = w^+ - w^-$. The dynamics of the probabilities (P_1, P_2) is described by a rate equation

$$\frac{d}{dt}\boldsymbol{P} = \mathcal{W}\boldsymbol{P}, \qquad\qquad \mathcal{W} = \begin{pmatrix} -w^+ & w^- \\ w^+ & -w^- \end{pmatrix} \tag{5.7}$$

Using the notation $S = P_2 - P_1$ for the probability difference, and recalling that $P_1 + P_2 = 1$, we get the equation

$$\frac{dS}{dt} \;=\; -\gamma S + u \tag{5.8}$$

which implies exponential relaxation towards the equilibrium value $S_{\rm eq} = u/\gamma$. If the transitions are induced by a bath of temperature T_B, then $S_{\rm eq}$ corresponds to equilibrium at temperature T_B. If we expose the system to so-called "work agent", say a sun that has infinite temperature, then the new rates are $w^\pm = w_B^\pm + w_A$, and the system will reach an equilibrium-like state that corresponds to a higher temperature. It is important to realize that the steady state solution features energy flow from the work agent via the system to the bath (and not the other way around). We say that the energy is dissipated.

======= [5.4] Three level/site system

The three-level system is the simplest setup for illustration of non-equilibrium thermodynamics. For example, it can be regarded as a model for a 3-level laser heat engine (see figure) or a mathematically equivalent rolling marble machine (see figure). The transitions are induced by a hot bath (T_H) and by a cold bath (T_C). In the first example photons can be either emitted or absorbed by a work agent ($T_A = \infty$). The second example is further discussed below. Either way the dynamics is generated by the matrix

$$\mathcal{W} \;=\; \begin{pmatrix} -\gamma_1 & w_C^- & w_A \\ w_C^+ & -\gamma_2 & w_H^- \\ w_A & w_H^+ & -\gamma_3 \end{pmatrix} \tag{5.9}$$

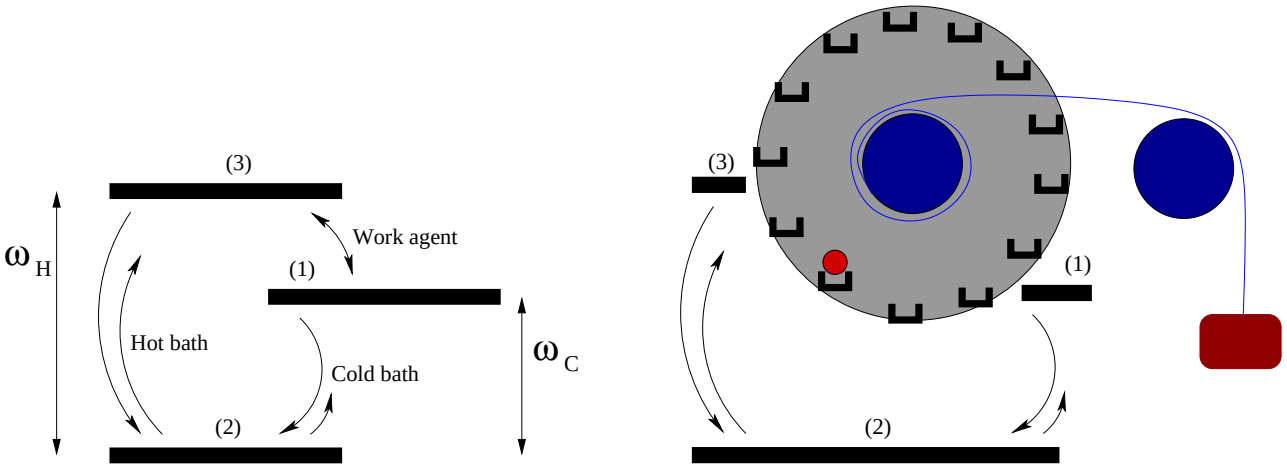
where $\gamma_1 = w_C^+ + w_A$, and $\gamma_2 = w_C^- + w_H^+$, and $\gamma_3 = w_H^- + w_A$. The affinity of the cycle is defined as

$$\Phi \;=\; \mathcal{E}_{1\rightsquigarrow 2} + \mathcal{E}_{2\rightsquigarrow 3} + \mathcal{E}_{3\rightsquigarrow 1} \;=\; \ln\left[\frac{w_{13}w_{32}w_{21}}{w_{12}w_{23}w_{31}}\right] \;=\; \frac{\omega_C}{T_C} - \frac{\omega_H}{T_H} \tag{5.10}$$

In order to have a working engine cycle we require $\Phi > 0$, which implies $(\omega_C/\omega_H) > (T_C/T_H)$. The efficiency of the engine is

$$\eta \;\equiv\; \frac{\omega_H - \omega_C}{\omega_H} \;<\; 1 - \frac{T_C}{T_H} \tag{5.11}$$

The limiting efficiency is the so-called Carnot efficiency. We can solve $\mathcal{W}\boldsymbol{P} = 0$ to find the probabilities (P_1, P_2, P_3) at steady state. Then we can find the probability current $I(\Phi) = (P_3 - P_1)w_A$ at steady state, and the power output of the engine $(\omega_H - \omega_C)I(\Phi)$.



Work agent.– The mechanical rolling marble machine possibly clarifies better the concept of *work agent*. Here the task of the engine is to pull up a weight. The hot bath induces, with some probability, a transition of the marble form position "2" to position "3". From there, with some probability, it gets into a car of the roller coaster wheel. Then is rolls (trapped in the car) to position "1". The wheel pulls up the weight. In order to maximize efficiency it is designed such that the potential energy of the whole system (including the weight) is the same at "3" and at "1". Consequently there is an equal probability to do the ride from "1" to "3". However, considering the full cycle, the condition $\Phi > 0$ ensures that the net work is positive.

═══ [5.5] The N site ring

The dynamics of a particle in an N site ring is generated by the matrix

$$\mathcal{W} = \begin{pmatrix} -\gamma & w^- & 0 & ... & w^+ \\ w^+ & -\gamma & w^- & 0 & ... \\ 0 & w^+ & -\gamma & w^- & ... \\ ... & ... & ... & ... & w^- \\ w^- & 0 & ... & w^+ & -\gamma \end{pmatrix} = -\gamma \mathbf{1} + w^+ \boldsymbol{D} + w^- \boldsymbol{D}^{-1} \tag{5.12}$$

We have assumed that all the anti-clockwise rates equal w^+, and that all the clockwise rates equal w^-. Accordingly $\gamma = w^+ + w^-$. The matrices $\boldsymbol{D}$ and $\boldsymbol{D}^{-1}$ generate anti-clockwise and clockwise displacements respectively. The drift velocity is

$$\frac{d}{dt}\langle x \rangle = [w^+ - w^-]a \tag{5.13}$$

where a is the lattice spacing. The proof is as follows:

$$\begin{aligned}\frac{d}{dt}\langle x\rangle &= \frac{d}{dt}\sum_n P_n x_n = \sum_n x_n \mathcal{W}_{nm} P_m = \sum_n x_n \left[w^+ P_{n-1} + w^- P_{n+1} - \gamma P_n\right] \\ &= \sum_n \left[w^+(x_{n+1}-x_n)P_n + w^-(x_{n-1}-x_n)P_n\right] = (w^+ - w^-)a \end{aligned} \tag{5.14}$$

Irrespective of drift, we have diffusion. Say that we start a distribution at $x = 0$. For simplicity let us assume that $w^- = w^+ = w$. We get that the rate of growth of the spreading is

$$\frac{d}{dt}\langle x^2\rangle = 2wa^2 \equiv 2D \tag{5.15}$$

More generally, if the drift velocity is non-zero, we can prove that

$$\mathrm{Var}(x) = \langle x^2\rangle - \langle x\rangle^2 = 2Dt \tag{5.16}$$

where $D = (1/2)[w^+ + w^-]a^2$ is called the diffusion coefficient.

[5.6] Relaxation modes

The steady state of the stochastic system is found from the equation $\mathcal{W}\boldsymbol{P}^{SS} = 0$. The relaxation modes are the eigenstates, namely $\mathcal{W}\psi = -\lambda\psi$, where $\{-\lambda\}$ are the eigenvalues. Note the sign convention, and note that the $\lambda = 0$ mode is the steady-state. An arbitrary initial state can be expanded in this basis, and consequently the solution of the rate equation is

$$\boldsymbol{P}(t) = e^{\mathcal{W}t}\boldsymbol{P}(0) = \boldsymbol{P}^{SS} + \sum_{\lambda\neq 0} C_\lambda\, e^{-\lambda t}\, \psi^{(\lambda)} \tag{5.17}$$

As an example we consider the N site ring. The eigenstates of $\mathcal{W}$ are the eigenstates of the displacement operator $\boldsymbol{D}$. Those eigenstates are called *momentum states*. Before we proceed note about notations: It is convenient to write the column representation ψ_n as a function, namely $\psi_n \equiv \psi(x_n)$. It is easily verified that the solution of $\boldsymbol{D}\psi = \tilde{\lambda}\psi$ are the functions $\psi(x) = e^{ikx}$. Operating on them with $\boldsymbol{D}$ we get $\boldsymbol{D}\psi(x) = e^{ik(x-a)}$. We deduce that the eigenvalues are $\tilde{\lambda} = e^{-ika}$. With standard normalization the momentum states are written as follows:

$$|k\rangle = \frac{1}{\sqrt{N}}\sum_n e^{ikx_n}\,|n\rangle \tag{5.18}$$

In the present context the momentum states have formal (non-physical) meaning. They are useful in order to write the solution of the rate equation as a sum of the $|k{=}0\rangle$ steady state and the $|k{\neq}0\rangle$ decaying relaxation modes. Namely,

$$P_n(t) = \sum_k C_k\, e^{-\lambda_k t}\; e^{ikx_n} \tag{5.19}$$

where

$$\lambda_k = \gamma - w^+ e^{-ika} - w^- e^{ika} = \gamma[1-\cos(ka)] + i\bar{v}\sin(ka) \tag{5.20}$$

We shall discuss further the k states in a later lecture.

Mathematics

[6] Algebra

The required background for this course: Complex numbers; Familiarity with elementary functions; and Linear algebra.

[6.1] Complex numbers

Real number x is the coordinate of a point along an axis, aka "the real axis". Complex number $z = (a, b)$ is the coordinate of a point in a the so called "complex plane". Common notation $i \equiv (0, 1)$, hence $z = a + ib$. Polar representation in terms of (r, φ). In multiplication the r-s are multiplied and the φ-s are added. Hence $i^2 = -1$. Complex conjugation $z^* = a - ib$, which means in polar coordinates $\varphi \mapsto -\varphi$. The absolute value is $r = |z| = \sqrt{z^* z} = \sqrt{a^2 + b^2}$. From Taylor expansion we get the Euler formula $e^{i\varphi} = \cos(\varphi) + i \sin(\varphi)$, hence

$$z = a + ib = r\cos(\varphi) + ir\sin(\varphi) = re^{i\varphi} \tag{6.1}$$

The N roots of the equation $z^N = 1$ are $z = e^{-i(2\pi/N)n}$, where n is an integer modulo N. The Fundamental theorem of algebra tells us that any polynomial of degree N has N roots, and can be factored accordingly. A quadratic equation with real coefficients might have two real root or two complex-conjugated roots.

[6.2] Differential equations

We encounter in this course two differential equations. One can be regrded as describing population dynamics, or stochastic relaxation, or the dynamics of an RC circuit:

$$\text{Initial condition:} \quad f(0) = f_0 \tag{6.2}$$

$$\text{Equation:} \quad \frac{df}{dt} = -\lambda f + c \tag{6.3}$$

$$\text{Solution:} \quad f(t) = \frac{c}{\lambda} + \left(f_0 - \frac{c}{\lambda}\right) e^{-\lambda t} \tag{6.4}$$

The other equation of interest describes harmonic motion:

$$\text{Equation:} \quad \frac{d^2 f}{dt^2} = -\omega^2 f \tag{6.5}$$

$$\text{Solution:} \quad f(t) = a\cos(\omega t) + b\sin(\omega t) = |A|\cos(\varphi - \omega t) = \Re\left[Ae^{-i\omega t}\right] \tag{6.6}$$

where the $A = a + ib = |A|e^{i\varphi}$ is the complex amplitude. The a and the b are determined by the initial "position" and the initial "velocity" respectively. Namely, $f(0) = a$ and $\dot{f}(0) = \omega b$.

[6.3] Linear algebra

In Euclidean geometry, three dimensional vectors can be written as:

$$\vec{u} = u_1\vec{e}_1 + u_2\vec{e}_2 + u_3\vec{e}_3 \tag{6.7}$$

Using Dirac notation we can write the same as:

$$|u\rangle \quad = \quad u_1|e_1\rangle + u_2|e_2\rangle + u_3|e_3\rangle \tag{6.8}$$

We say that the vector has the representation:

$$|u\rangle \quad \mapsto \quad u_i \; = \; \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \tag{6.9}$$

The operation of a linear operator A is written as $|v\rangle = A|u\rangle$ which is represented by:

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \; = \; \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \tag{6.10}$$

or shortly as $v_i = A_{ij}u_j$. Thus a linear operator is represented by a matrix:

$$A \quad \mapsto \quad A_{ij} \; = \; \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \tag{6.11}$$

═══ [6.4] Inner product

The length of a vector is a generalization of the "absolute value" of complex number. Namely

$$|u|^2 \quad = \quad \sum_j |u_j|^2 \quad = \quad \sum_j u_j^* u_j \tag{6.12}$$

More generally we define an inner product

$$\langle u|v\rangle \quad = \quad \sum_j u_j^* v_j \quad = \quad u_1^* v_1 + u_2^* v_2 + u_3^* v_3 + ... \tag{6.13}$$

Accordingly $|u| = \sqrt{\langle u|u\rangle}$. A basis is called orthonormal if

$$\langle e_i|e_j\rangle \quad = \quad \delta_{i,j} \tag{6.14}$$

It follows that the representation of a given vector can be calculated by pojecting it on the basis-vectors:

$$u_j \quad = \quad \langle e_j|u\rangle \tag{6.15}$$

For the matrix elements it follows that

$$A_{ij} \quad = \quad \langle e_i|A|e_j\rangle \tag{6.16}$$

[6.5] Hilbert space

Later we are going to say that the pure-states of a system form a linear vector space with an inner-product. The inner-product tells us whether the states are orthogonal. We call such space *Hilbert space*. We shall not use the term "vector", instead we shall use the term "state". Vectors that differ by normalization or by phase factor (aka gauge) define the same state. More elegant phrasing is to say that *pure states* of the system have one-to-one correspondence with *projectors* over Hilbert space.

[6.6] Completeness of the basis

In Dirac notation the expansion of a vector is written as:

$$|u\rangle \;=\; |e_1\rangle\langle e_1|u\rangle + |e_2\rangle\langle e_2|u\rangle + |e_3\rangle\langle e_3|u\rangle \tag{6.17}$$

which implies

$$\mathbf{1} \;=\; |e_1\rangle\langle e_1| + |e_2\rangle\langle e_2| + |e_3\rangle\langle e_3| \tag{6.18}$$

Above $\mathbf{1} \mapsto \delta_{ij}$ stands for the identity operator, and $P^j = |e_j\rangle\langle e_j|$ are called "projector operators",

$$\mathbf{1} \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P^1 \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad P^2 \mapsto \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad P^3 \mapsto \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{6.19}$$

Now we can define the "completeness of the basis" as the requirement

$$\sum_j P^j \;=\; \sum_j |e_j\rangle\langle e_j| \;=\; \mathbf{1} \tag{6.20}$$

From the completeness of the basis it follows e.g. that for any operator

$$A \;=\; \left[\sum_i P^i\right] A \left[\sum_j P^j\right] \;=\; \sum_{i,j} |e_i\rangle\langle e_i|A|e_j\rangle\langle e_j| \;=\; \sum_{i,j} |e_i\rangle A_{ij}\langle e_j| \tag{6.21}$$

[6.7] Operators

In what follows we are interested in "normal" operators that are diagonal in some orthonormal basis. Say that we have an operator A. By definition, if it is normal, there exists an orthonormal basis $\{|a\rangle\}$ such that A is diagonal. Hence we write

$$A \;=\; \sum_a |a\rangle a\langle a| \;=\; \sum_a a P^a \tag{6.22}$$

In matrix representation it means:

$$\begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} \;=\; a_1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + a_3 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{6.23}$$

It is useful to define what is meant by $\hat{B} = f(\hat{A})$ where $f()$ is an arbitrary function. Assuming that $\hat{A} = \sum |a\rangle a\langle a|$, it follows by definition that $\hat{B} = \sum |a\rangle f(a)\langle a|$. Another useful rule to remember is that if $A|k\rangle = B|k\rangle$ for some complete basis k, then it follows by linearity that $A|\psi\rangle = B|\psi\rangle$ for any vector, and therefore $A = B$.

With any operator A, we can associate an "adjoint operator" $A^\dagger$. By definition it is an operator that satisfies the following relation:

$$\langle u|Av\rangle = \langle A^\dagger u|v\rangle \tag{6.24}$$

If we substitute the basis vectors in the above relation we get the equivalent matrix-style definition

$$(A^\dagger)_{ij} \quad = \quad A^*_{ji} \tag{6.25}$$

If A is normal then it is diagonal in some orthonormal basis, and then also $A^\dagger$ is diagonal in the same basis. It follows that a normal operator has to satisfy the *necessary* condition $A^\dagger A = AA^\dagger$. As we show below this is also a *sufficient* condition for "normality".

We first consider Hermitian operators, and show that they are "normal". By definition they satisfy $A^\dagger = A$. If we write this relation in the eigenstate basis we deduce after one line of algebra that $(a^* - b)\langle a|b\rangle = 0$, where a and b are any two eigenvalues. If follows (considering $a = b$) that the eigenvalues are *real*, and furthermore (considering $a \neq b$) that eigenvectors that are associate with different eigenvalues are orthogonal. This is called the *spectral theorem*: one can find an orthonormal basis in which A is diagonal.

We now consider a general operator Q. Always we can write it as

$$Q = A + iB, \qquad \text{with } A = \frac{1}{2}(Q + Q^\dagger), \qquad \text{and } B = \frac{1}{2i}(Q - Q^\dagger) \tag{6.26}$$

One observes that A and B are Hermitian operators. It is easily verified that $Q^\dagger Q = QQ^\dagger$ iff $AB = BA$. It follows that there is an orthonormal basis in which both A and B are diagonal, and therefore Q is a normal operator.

We see that an operator is normal iff it satisfies the commutation $Q^\dagger Q = QQ^\dagger$ and iff it can be written as a function $f(H)$ of an Hermitian operator H. We can regard any H with non-degenerate spectrum as providing a specification of a basis, and hence any other operator that is diagonal in that basis can be expressed as a function of this H.

Of particular interest are unitary operators. By definition they satisfy $U^\dagger U = \mathbf{1}$, and hence they are "normal" and can be diagonalized in an orthonormal basis. Hence their eigenvalues satisfy $\lambda_r^* \lambda_r = 1$, which means that they can be written as:

$$U \quad = \quad \sum_r |r\rangle \mathrm{e}^{i\varphi_r} \langle r| \quad = \quad \mathrm{e}^{iH} \tag{6.27}$$

where H is Hermitian. This is an example for the general statement that any normal operator can be written as a function of some Hermitian operator H.

[6.8] Generators

Later we are going to discuss unitary operations that depend on some parameter. For example: $\boldsymbol{D}(a)$ is distance a displacement; $\boldsymbol{R}(\Phi)$ is angle Φ rotation; and $\boldsymbol{U}(t)$ is evolution during time t. Note that $\boldsymbol{D}(0) = \boldsymbol{R}(0) = \boldsymbol{U}(0) = \mathbf{1}$. We shall define generators for those operations. For example, for evolution we define $\boldsymbol{U}(t) = \exp(-it\boldsymbol{H})$.

Proof: We assume that the operation, say the evolution, has the so-called group property

$$U(t_2 + t_1) \quad = \quad U(t_2)\, U(t_1) \tag{6.28}$$

It follows that

$$U(t) \quad = \quad [U(t/N)]^N \tag{6.29}$$

where N can be as large as we want. The evolution during an infinitesimal time interval can be written as:

$$U(dt) \quad = \quad \hat{1} + Gdt + \mathcal{O}(dt^2) \quad \equiv \quad \hat{1} - idt\mathcal{H} + \mathcal{O}(dt^2) \tag{6.30}$$

The first equality is a Taylor expansion, or one may say that G is the derivative of $U(t)$. Loosely speaking G is the "evolution per unit of time". From $U^\dagger U = 1$ it follows that $G^\dagger = -G$. The physics community does not like anti-hermitian operators, and therefore we define $\mathcal{H} \equiv iG$, that satisfies $\mathcal{H}^\dagger = \mathcal{H}$. By multiplying many infinitesimal time steps we get:

$$\hat{U} \quad = \quad (1 - idt\mathcal{H}) \cdots (1 - idt\mathcal{H})(1 - idt\mathcal{H}) \quad = \quad \left(1 - i\frac{t}{N}\mathcal{H}\right)^N \quad = \quad \mathrm{e}^{-it\mathcal{H}} \tag{6.31}$$

In the last step we have taken the $N \to \infty$ limit, and used the definition of the exponential function, Namely, $\exp(t) = \lim(1 + t/N)^N$, that follows from the assumed multiplicative property $\exp(t_1)\exp(t_2) = \exp(t_1 + t_2)$.

[6.9] Conventions regarding notations

In Mathematica there is a clear distinction between dummy indexes and fixed values. For example $f(x\_) = 8$ means that $f(x) = 8$ for any x, hence x is a dummy index. But if $x = 4$ then $f(x) = 8$ means that only one element of the vector $f(x)$ is specified. Unfortunately in the printed mathematical literature there are no clear conventions. However the tradition is to use notations such as $f(x)$ and $f(x')$ where x and x' are dummy indexes, while $f(x_0)$ and $f(x_1)$ where x_0 and x_1 are fixed values. Thus

$$A_{ij} \quad = \quad \begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix} \tag{6.32}$$
$$A_{i_0 j_0} \quad = \quad 5 \qquad \text{for } i_0 = 2 \text{ and } j_0 = 1$$

Another typical example is

$$T_{x,k} \quad = \quad \langle x|k\rangle \quad = \quad \text{matrix} \tag{6.33}$$
$$\Psi(x) \quad = \quad \langle x|k_0\rangle \quad = \quad \text{column} \tag{6.34}$$

In the first equality we regard $\langle x|k\rangle$ as a matrix: it is the transformation matrix form the position to the momentum basis. In the second equality we regard the same object (with fixed k_0) as a column, or as a "wave-function".

[7] Translations and momentum

======= [7.1] N site system

A site is a location where a particle can be positioned. If we have $N = 5$ sites it means that we have a 5-dimensional Hilbert space of quantum states. Later we shall assume that the particle can "jump" between sites. For mathematical reasons it is conveneint to assume torus topology. This means that the next site after $x = 5$ is $x = 1$. This is also called periodic boundary conditions.

The standard basis is the position basis. For example: $|x\rangle$ with $x = 1, 2, 3, 4, 5$. So we can define the position operator as follows:

$$\hat{x}|x\rangle = x|x\rangle \tag{7.1}$$

In this example we get:

$$\hat{x} \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix} \tag{7.2}$$

The operation of this operator on a state vector is for example:

$$\begin{aligned} &|\psi\rangle = 7|3\rangle + 5|2\rangle \\ &\hat{x}|\psi\rangle = 21|3\rangle + 10|2\rangle \end{aligned} \tag{7.3}$$

======= [7.2] Translation operators

The one-step translation operator is defined as follows:

$$\hat{D}|x\rangle = |x + 1\rangle \tag{7.4}$$

For example:

$$D \mapsto \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \tag{7.5}$$

and hence $D|1\rangle = |2\rangle$ and $D|2\rangle = |3\rangle$ and $D|5\rangle = |1\rangle$. Let us consider the superposition:

$$|\psi\rangle = \frac{1}{\sqrt{5}}[|1\rangle + |2\rangle + |3\rangle + |4\rangle + |5\rangle] \tag{7.6}$$

It is clear that $D|\psi\rangle = |\psi\rangle$. This means that ψ is an eigenstate of the translation operator (with eigenvalue e^{i0}). The translation operator has other eigenstates that we will discuss in the next section.

[7.3] Momentum states

The momentum states are defined as follows:

$$|k\rangle \to \frac{1}{\sqrt{N}} e^{ikx} \tag{7.7}$$
$$k = \frac{2\pi}{N} n, \quad n = \text{integer} \quad \mod (N)$$

In the previous section we have encountered the $k = 0$ momentum state. In Dirac notation this is written as:

$$|k\rangle = \sum_x \frac{1}{\sqrt{N}} e^{ikx} |x\rangle \tag{7.8}$$

or equivalently as:

$$\langle x|k\rangle = \frac{1}{\sqrt{N}} e^{ikx} \tag{7.9}$$

while in old fashioned notation it is written as:

$$\psi_x^k = \langle x|k\rangle \tag{7.10}$$

where the upper index k identifies the state, and the lower index x is the representation index. Note that if x were continuous then it would be written as $\psi^k(x)$.

The k states are eigenstates of the translation operator. This can be proved as follows:

$$D|k\rangle = \sum_x D|x\rangle\langle x|k\rangle = \sum_x |x+1\rangle \frac{1}{\sqrt{N}} e^{ikx} = \sum_{x'} |x'\rangle \frac{1}{\sqrt{N}} e^{ik(x'-1)} = e^{-ik} \sum_{x'} |x'\rangle \frac{1}{\sqrt{N}} e^{ikx'} = e^{-ik}|k\rangle \tag{7.11}$$

Hence we get the result:

$$D|k\rangle = e^{-ik}|k\rangle \tag{7.12}$$

and conclude that $|k\rangle$ is an eigenstate of $\hat{D}$ with an eigenvalue e^{-ik}. Note that the number of independent eigenstates is N. For exmaple for a 5-site system we have $e^{ik_6} = e^{ik_1}$.

[7.4] Momentum operator

The momentum operator is defined as follows:

$$\hat{p}|k\rangle \equiv k|k\rangle \tag{7.13}$$

From the relation $\hat{D}|k\rangle = \mathrm{e}^{-ik}|k\rangle$ it follows that $\hat{D}|k\rangle = \mathrm{e}^{-i\hat{p}}|k\rangle$. Therefore we deduce the operator identity:

$$\hat{D} \;=\; \mathrm{e}^{-i\hat{p}} \tag{7.14}$$

We can also define 2-step, 3-step, and a-step translation operators as follows:

$$\begin{array}{rcccl} \hat{D}(2) & = & (\hat{D})^2 & = & \mathrm{e}^{-i2\hat{p}} \\ \hat{D}(3) & = & (\hat{D})^3 & = & \mathrm{e}^{-i3\hat{p}} \\ \hat{D}(a) & = & (\hat{D})^a & = & \mathrm{e}^{-ia\hat{p}} \end{array} \tag{7.15}$$

======= [7.5] The algebra of the generator

The following identity reflects the idea that $\hat{p}$ is the generator of a displacements:

$$[\hat{x}, \hat{D}(a)] \;=\; a\hat{D} \tag{7.16}$$

In order to prove it, let us see that the identity is correct for any basis state, and hence it is correct for any state.

$$\begin{array}{l} [\hat{x}, D]\,|x_0\rangle \;=\; (\hat{x}D - D\hat{x})\,|x_0\rangle \;=\; \hat{x}D\,|x_0\rangle - D\hat{x}\,|x_0\rangle \;=\; \hat{x}\,|x_0{+}a\rangle - x_0 D\,|x_0\rangle \\ \qquad =\; (x_0{+}a)\,|x_0{+}a\rangle - x_0\,|x_0{+}a\rangle \;=\; a\,|x_0{+}a\rangle \;=\; aD\,|x_0\rangle \end{array}$$

The infinitesimal version of this relation is

$$[\hat{x}, \hat{p}] \;=\; i \tag{7.17}$$

And a more general statement that can be proved for any function $g(p)$ is

$$[\hat{x}, g(\hat{p})] \;=\; ig'(p) \tag{7.18}$$

The dual relation for any function $f(x)$ is

$$[f(\hat{x}), \hat{p}] \;=\; if'(x) \tag{7.19}$$

[8] Rotations and Spin

[8.1] Groups

A group is a set of elements with a binary operation:

- The operation is defined by a multiplication table for $\tau * \tau'$.
- There is a unique identity element $\mathbf{1}$.
- Every element has an inverse element such that $\tau\tau^{-1} = \mathbf{1}$
- Associativity: $\tau * (\tau' * \tau'') = (\tau * \tau') * \tau''$

Commutativity does not have to be obeyed: this means that in general $\tau * \tau' \neq \tau' * \tau$. In practice, a group that has a finite number of elements, can be defined via a *multiplication table*. But this is strategy is not practical if we are dealing with a group that has infinite number of elements. Such is the group of rotations SO(3).

For a Lie group, we assume that each element τ of the group is labeled by a set of parameters that serve as coordinates. Therefore we can regard the set of elements as a set of points that form a *manifold*. By convention we label the identity $\mathbf{1}$ as $\tau = (0, 0, ..., 0)$, and refer to it as the origin of the manifold. For SO(3) the standard labeling of rotations is $\vec{\Phi} = (\Phi_x, \Phi_y, \Phi_z)$, and the manifold is a ball of radius π.

In order to provide a tangible definition for the group operation, each element of the group is regarded as a *transformation* over some space:

$$\text{element of the group} \quad = \quad \boldsymbol{\tau} = (\tau_1, \tau_2, ..., \tau_d) \tag{8.1}$$

$$\text{corresponding transformation} \quad = \quad U(\boldsymbol{\tau}) \tag{8.2}$$

$$\text{the group property} \quad = \quad U(\boldsymbol{\tau'})U(\boldsymbol{\tau''}) = U(\boldsymbol{\tau'} * \boldsymbol{\tau''}) \tag{8.3}$$

The association $\boldsymbol{\tau} \mapsto U(\boldsymbol{\tau})$ is called a *realization*. We use the notations $\tau * \tau * ... * \tau \equiv n\tau$, where n is the number of repetitions. Accordingly $U(0) = \mathbf{1}$, and $U(n\tau) = U(\tau)U(\tau)...U(\tau) = U(\tau)^n$. The crucial observation for a Lie group is that we can generate large transformations from infinitesimal transformations. Thus it is meaningful to define a *ray* of elements as $U(\alpha\tau)$ where α is a real number. For SO(3) a ray of elements includes all the rotations around the same axis.

[8.2] Realizations and Representations

As defined above, a realization means that we regard each element of the group as an operation over a space. We treat the elements as transformations. Below we discuss the possibility of finding a realization that consists of *linear* transformations. Such transformations are represented by matrices. Consequenctly, with each element $\boldsymbol{\tau}$ of the group, we associate a matrix $U_{ij}(\boldsymbol{\tau})$. The "multiplication table" of the matrices is required to be be in one-to-one correspondence to the multiplication table of the elements of the group. Below we "soften" this requirement, being satisfied with having a "multiplication table" that is the same "up to a phase factor":

$$U(\boldsymbol{\tau'} * \boldsymbol{\tau''}) \;=\; \mathrm{e}^{i(\text{phase})} U(\boldsymbol{\tau'})U(\boldsymbol{\tau''}) \tag{8.4}$$

It is natural to consider realizations that consist of orthogonal transformations over a real space, or unitary transformations over a complex space. Any realization using a linear transformations provides a "representation" once a basis is chosen: linear transformations are represented by matrices.

Finding a "representation" for a group is very useful: the operative meaning of "multiplying group elements" becomes "multiplying matrices". This means that we can deal with groups using linear algebra tools.

[8.3] The rotation group SO(3)

The rotation group $SO(3)$ is a non-commutative group. That means that the order of rotations is important. Its defining representation consists of all the real unitary matrices (aka Orthogonal matrices) that have determinant that equals unity. Such matrices rotate vectors in the Euclidean 3D space, without changing their length. The positivity of the determinant implies that we excluded reflections (mirror operations) from the group.

An element of SO(3) is parameterized by

$$\vec{\Phi} \equiv \Phi\vec{n} = (\Phi_x, \Phi_y, \Phi_z) \tag{8.5}$$

where Φ is the angle of rotation, and $\vec{n}$ is a unit vector that indicates the axis of rotation. In spherical coordinates

$$\vec{n} = (\sin\theta\cos\varphi,\ \sin\theta\sin\varphi,\ \cos\theta) \tag{8.6}$$

It is trivial to write SO(3) matrix that represent rotation around the Z axis in the standard basis:

$$R(\Phi\vec{e}_z) = \begin{pmatrix} \cos(\Phi) & -\sin(\Phi) & 0 \\ \sin(\Phi) & \cos(\Phi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{8.7}$$

For a small rotation:

$$R(\delta\Phi\vec{e}_z) = \begin{pmatrix} 1 & -\delta\Phi & 0 \\ \delta\Phi & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \hat{1} + \delta\Phi \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \hat{1} - i\delta\Phi M_z \tag{8.8}$$

Thus, the generator of rotations around the Z axis is

$$M_z = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{8.9}$$

We can find the other generators in the same way:

$$M_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \qquad M_y = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \tag{8.10}$$

The way to construct a general rotation matrix (proof below) is

$$R(\vec{\Phi}) = \mathrm{e}^{-i\Phi M_n} \tag{8.11}$$

where the generator of rotation around the $\vec{n}$ axis is

$$M_n = n_x M_x + n_y M_y + n_z M_z \equiv \vec{n}\cdot\vec{M} \tag{8.12}$$

It is important to realize that contrary to translation we are dealing here with a non-commutative group, and therefore in general

$$R(\vec{\Phi}) \equiv R(\vec{\Phi}'')R(\vec{\Phi}') \neq R(\vec{\Phi}')R(\vec{\Phi}'') \neq R(\vec{\Phi}'+\vec{\Phi}'') \tag{8.13}$$

Given Φ' and Φ'' there is no trivial way to find Φ. For example if we rotate a body 90deg around the X axis, and after that 90deg around the Y axis, the outcome is *not* a rotation around the the $\vec{n} \propto (1,1,0)$ axis, but rather a 120deg rotation around the $\vec{n} \propto (1,1,-1)$ axis... The way to make such a deduction is to calculate the matrices $R(\vec{\Phi}'')$ and $R(\vec{\Phi}')$, then to multiply them to get equivalent rotation matrix R, and then to extract $\vec{\Phi}$ by writing the result in terms of the generators as $R = \mathrm{e}^{-i\Phi M_n}$.

═══ [8.4] Proof of $M_n = \vec{n} \cdot \vec{M}$

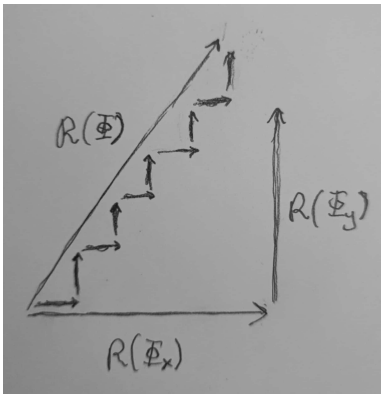
The following proof is based on the observation that given matrices A and B, the product $e^A e^B$ is not the same as e^{A+B}, because the multiplication of the matrices is not commutative. On the other hand for arbitrarily small ϵ the equality $e^{\epsilon A}e^{\epsilon B} = e^{\epsilon(A+B)}$ is arbitrarily accurate, as implied by first-order Taylor expansion of the exponent.

Given that we know how to construct rotation matrices say around the X and the Y axes, we would like to have a recipe to construct a general rotation matrix around an arbitrary axis $\vec{n}$. For that we need an expression for the generator M_n. We would like to prove that $M_n = \vec{n} \cdot \vec{M}$.

Without loss of generality, let us assume that $\vec{\Phi} = (\Phi_x, \Phi_y, 0)$. As already explained, the naive prescription for construction a general rotation matrix is wrong. Namely, in general, the decomposition $R(\vec{\Phi}) = R(\Phi_x \vec{e}_y)R(\Phi_y \vec{e}_x)$ is not valid. On the other hand, for a very small angle $\Phi := \Phi/N$, where N is arbitrarily large, this equality is arbitrarily accurate. Equipped with this observation we deduce the following

$$\begin{aligned}
R(\vec{\Phi}) &= \Big[R(\vec{\Phi}/N)\Big]^N && (8.14)\\
&= [R(\Phi_y/N)\ R(\Phi_x/N)]^N && (8.15)\\
&= [\exp(-i(\Phi_y/N)M_y)\ \exp(-i(\Phi_x/N)M_x)]^N && (8.16)\\
&= [\exp(-i(\Phi_y/N)M_y - i(\Phi_x/N)M_x)]^N && (8.17)\\
&= \exp(-i(\Phi_x M_x + \Phi_y M_y)) && (8.18)\\
&= \exp\left(-i\vec{\Phi}\cdot\vec{M}\right) = \exp(-i\Phi M_n) && (8.19)
\end{aligned}$$

The logic of the proof is illustrated using the following schematic figure:



[8.5] The Lie Algebra

In general, the result of $e^A e^B$ is not e^{A+B}. The result depends on the commutator of the matrices

$$[A,B] \equiv AB - BA \tag{8.20}$$

Note that $[B,A] = -[A,B]$. In the present context the multiplication table of the rotation-matrices is determined by the commutators of the generators. This is the so-called Lie Algebra of the group:

$$[M_x, M_y] = iM_z \tag{8.21}$$
$$[M_y, M_z] = iM_x \tag{8.22}$$
$$[M_z, M_x] = iM_y \tag{8.23}$$

These commutation relations serve as the DNA of the rotation group. Namely, it is the starting point for the construction of other representations of the group, as illustrated in the next section.

[8.6] The dim=2 representation

We have defined the rotation group by the Euclidean realization over 3D space. Obviously, this representation can be used to make calculations ("to multiply rotations"). The advantage is that it is intuitive, and there is no need for complex numbers. The disadvantage is that they are 3×3 matrices with inconvenient algebraic properties, so a calculation could take hours. It would be convenient if we could "multiply rotations" with simple 2×2 matrices. In other words, we are interested in a dim=2 representation of the rotation group. The mission is to find three simple 2×2 matrices that satisfy the same Lie Algebra:

$$[S_x, S_y] = iS_z \tag{8.24}$$
$$[S_y, S_z] = iS_x \tag{8.25}$$
$$[S_z, S_x] = iS_y \tag{8.26}$$

There is a systematic approach for building all the representations of the rotation group. In the present lecture, we simply find the requested representation by guessing. First we recall the definition of the Pauli matrices

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{8.27}$$

Any 2×2 matrix can be written as a linear combination of these 4 matrices. In particular all the linear combinations $\sigma_n = \vec{n} \cdot \vec{\sigma}$, where $\vec{n}$ is a unit vector, have the property $\sigma_n^2 = 1$, and therefore have the eigenvalues ± 1, which means that they are all *similar* to σ_z. We also note that multiplication of different Pauli matrices is very simple, namely $\sigma_x \sigma_y = i\sigma_z$, and similarly for all the cyclic permutations. Note also that for anti-cyclic multiplication we have $\sigma_y \sigma_x = -\sigma_x \sigma_y$. It follows that the matrices

$$S_x = \frac{1}{2}\sigma_x, \quad S_y = \frac{1}{2}\sigma_y, \quad S_z = \frac{1}{2}\sigma_z \tag{8.28}$$

satisfy the same commutation relations as those of (M_x, M_y, M_z). Therefore we can use them to generate a dim=2 representation of the rotation group. We construct the rotation matrices using the formula:

$$R(\vec{\Phi}) = \mathrm{e}^{-i\vec{\Phi}\cdot\vec{S}} \tag{8.29}$$

The matrices that are generated necessarily satisfy the group multiplication table.

═══ [8.7] The calculation of rotation matrices

The calculation of a 2×2 rotation matrix is extremely simple. All the even powers of a given Pauli matrix are equal to the identity matrix, while all the odd powers are equal to the original matrix. From this (using Taylor expansion and separating into two partial sums), we get the result:

$$R(\vec{\Phi}) \;=\; R(\Phi\vec{n}) \;=\; \mathrm{e}^{-i\Phi S_n} \;=\; \mathrm{e}^{-i(\Phi/2)\sigma_n} \;=\; \cos(\Phi/2)\hat{1} - i\sin(\Phi/2)\sigma_n \tag{8.30}$$

where $\sigma_n = \vec{n}\cdot\vec{\sigma}$, and $S_n = (1/2)\sigma_n$ is the generator of a rotation around the $\vec{n}$ axis. Of particular interest are rotations around the Z axis and around the Y axis:

$$R(\Phi\vec{e}_z) \;=\; \mathrm{e}^{-i\Phi S_z} \;=\; \begin{pmatrix} \mathrm{e}^{-i\Phi/2} & 0 \\ 0 & \mathrm{e}^{i\Phi/2} \end{pmatrix} \tag{8.31}$$

$$R(\Phi\vec{e}_y) \;=\; \mathrm{e}^{-i\Phi S_y} \;=\; \begin{pmatrix} \cos(\Phi/2) & -\sin(\Phi/2) \\ \sin(\Phi/2) & \cos(\Phi/2) \end{pmatrix} \tag{8.32}$$

The analogous formula for constructing a 3×3 rotation matrix is somewhat more complicated:

$$R(\vec{\Phi}) \;=\; R(\Phi\vec{n}) \;=\; \mathrm{e}^{-i\Phi M_n} \;=\; 1 + (\cos(\Phi) - 1)M_n^2 - i\sin(\Phi)M_n \tag{8.33}$$

where $M_n = \vec{n}\cdot\vec{M}$ is the generator of a rotation around the $\vec{n}$ axis. The proof is based on a Taylor expansion. We notice that $M_z^3 = M_z$, from this it follows that for all the odd powers $M_z^k = M_z$, while for all the even powers $M_z^k = M_z^2$ where $k > 0$. The same properties apply to any M_n, because all the rotations are "similar" (similarity transformation represents a change of basis).

═══ [8.8] An example for multiplication of rotations

Let us make a 90^o rotation $R(90^0 e_z)$ around the Z axis, followed by a 90^o rotation $R(90^0 e_y)$ around the Y axis. We would like to know what this sequence gives. Using the dim=3 Euclidean representation requires multiplication of 3×3 matrices, so we prefer to do the calculation with the simpler dim=2 representation.

$$\begin{aligned} R(\Phi) &= \cos(\Phi/2)\hat{1} - i\sin(\Phi/2)\sigma_n \\ R(90^0 e_z) &= \frac{1}{\sqrt{2}}(\hat{1} - i\sigma_z) \\ R(90^0 e_y) &= \frac{1}{\sqrt{2}}(\hat{1} - i\sigma_y) \end{aligned} \tag{8.34}$$

Hence

$$R \;=\; R(90^0 e_y)R(90^0 e_z) \;=\; \frac{1}{2}(1 - i\sigma_x - i\sigma_y - i\sigma_z) \tag{8.35}$$

where we have used the fact that $\sigma_y\sigma_z = i\sigma_x$. We can write this result as:

$$R = \cos\frac{120^o}{2} - i\sin\frac{120^o}{2}\vec{n}\cdot\vec{\sigma} \tag{8.36}$$

where $n = \frac{1}{\sqrt{3}}(1,1,1)$. This defines the equivalent rotation which is obtained by combining the two 90^o rotations.

═══ [8.9] Polarization states of Spin $1/2$

We now discuss the physical interpretation of the "states" that the $s = 1/2$ matrices rotate. Any state of "spin $1/2$" is represented by a vector with two complex number. That means we have 4 parameters. After gauge and normalization, we are left with 2 physical parameters which can be associated with the polarization direction (θ, φ). Thus it makes sense to represent the state of spin $1/2$ by an arrow that points to some direction in space.

The eigenstates of S_z do not change when we rotate them around the Z axis (aside from a phase factor). Therefore the following interpretation comes to mind:

$$\left|m = +\frac{1}{2}\right\rangle \;=\; |\vec{z}\rangle \;=\; |\uparrow\rangle \;\mapsto\; \begin{pmatrix}1\\0\end{pmatrix} \tag{8.37}$$

$$\left|m = -\frac{1}{2}\right\rangle \;=\; |\overleftarrow{z}\rangle \;=\; |\downarrow\rangle \;\mapsto\; \begin{pmatrix}0\\1\end{pmatrix} \tag{8.38}$$

This interpretation is confirmed by rotating the "up" state by 180 degrees, and getting the "down" state.

$$R \;=\; \mathrm{e}^{-i\pi S_y} \;=\; \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} \tag{8.39}$$

We see that:

$$\begin{pmatrix}1\\0\end{pmatrix} \;\rightsquigarrow\; 180^0 \;\rightsquigarrow\; \begin{pmatrix}0\\1\end{pmatrix} \;\rightsquigarrow\; 180^0 \;\rightsquigarrow\; -\begin{pmatrix}1\\0\end{pmatrix} \tag{8.40}$$

With two rotations of 180^o we get back the "up" state, with a minus sign. Optionally one observes that

$$\mathrm{e}^{-i2\pi S_z} \;=\; \mathrm{e}^{-i\pi\sigma_z} \;=\; -1 \tag{8.41}$$

and hence by similarity this holds for any 2π rotation. We see that the representation that we found is not a one-to-one representation of the rotation group. It does not obey the multiplication table in a one-to-one fashion! In fact, we have found a representation of $SU(2)$ and not $SO(3)$. The minus sign has a physical significance. In a two slit experiment it is possible to turn destructive interference into constructive interference by placing a magnetic field in one of the paths. The magnetic field rotates the spin of the electrons. If we induce 360^o rotation, then the relative phase of the interference change sign, and hence constructive interference becomes destructive and vice versa. The relative phase is important! Therefore, we must not ignore the minus sign.

It is important to emphasize that the physical degree of freedom that is called "spin $1/2$" cannot be visualized as a arising from the spinning of small rigid body around some axis like a top. If it were possible, then we could say that the spin can be described by a wave function. In this case, if we would rotate it by 360^o we would get the same state, with the same sign. But in the representation we are discussing we get minus the same state. That is in contradiction with the definition of a (wave) function as a single valued object.

We can get from the "up" state all the other possible states merely by using the appropriate rotation matrix. In particular we can get any spin polarization state by combining a rotation round the Y axis and a rotation round

the Z axis. The result is:

$$|\vec{n}_{\theta,\varphi}\rangle \quad = \quad R(\varphi)R(\theta)|\uparrow\rangle \quad = \quad \mathrm{e}^{-i\varphi S_z}\mathrm{e}^{-i\theta S_y}|\uparrow\rangle \quad \mapsto \quad \begin{pmatrix} \mathrm{e}^{-i\varphi/2}\cos(\theta/2) \\ \mathrm{e}^{i\varphi/2}\sin(\theta/2) \end{pmatrix} \tag{8.42}$$

that can be written in Dirac notations as

$$|\vec{n}_{\theta,\varphi}\rangle \quad = \quad \mathrm{e}^{-i\varphi/2}\cos(\theta/2)\,|\uparrow\rangle + \mathrm{e}^{i\varphi/2}\sin(\theta/2)\,|\downarrow\rangle \tag{8.43}$$

or in different gauge as

$$|\vec{n}_{\theta,\varphi}\rangle \quad = \quad \cos(\theta/2)\,|\uparrow\rangle + \mathrm{e}^{i\varphi}\sin(\theta/2)\,|\downarrow\rangle \tag{8.44}$$

Quantum Mechanics

[9] The evolution of quantum mechanical systems

[9.1] The evolution operator and the Hamiltonian

Consider a particle in an N site system. If we redefine the first basis state as $|\tilde{1}\rangle = -8|1\rangle$ we get a non-orthonormal basis, and therefore not convenient for practical mathematical calculations. *Explanation:* if we represent the state $|\psi\rangle$ as a linear combination of normalized basis vectors $|\psi\rangle = \sum_j \psi_j |j\rangle$, then we can find the coefficients of the combination by using the formula $\psi_i = \langle i|\psi\rangle$.

Even if we decide to work with a set of orthonormal states, still there is some freedom left which is called "gauge freedom" or "phase freedom": multiplying a state with a phase factor does not imply a different state, ρ stays the same, and hence all physical expectation values remain the same too. *Example:* The spin states $|\uparrow\rangle$ and $\mathrm{e}^{\frac{\pi}{8}i}|\uparrow\rangle$ are represented by the same polarization vector, or optionally by the same ρ.

From the superposition principle, and from the above remarks regarding the normalization, it follows that the evolution in quantum mechanics is described by a unitary operator:

$$\begin{aligned} |\psi^{t=0}\rangle &\rightarrow |\psi^t\rangle \\ |\psi^t\rangle &= \boldsymbol{\mathcal{U}}|\psi^{t=0}\rangle \end{aligned} \tag{9.1}$$

It is assumed here that the system is *isolated*. In order to simplify the discussion below we further assume that the external fields are constant in time. In such a case, the evolution operator has the group property $\boldsymbol{\mathcal{U}}(t_2+t_1) = \boldsymbol{\mathcal{U}}(t_2)\,\boldsymbol{\mathcal{U}}(t_1)$. It follows (see section 5.8) that the evolution operator can be written as

$$\boldsymbol{\mathcal{U}}(t) = \mathrm{e}^{-it\mathcal{H}} \tag{9.2}$$

where $\mathcal{H}$, the generator of the evolution, is called *Hamiltonian*.

[9.2] The Schrödinger Equation

Consider the evolution of a pure state $\psi^t = \boldsymbol{\mathcal{U}}\psi^{t=0}$. From the rerlation $\psi^{t+dt} = (1 - idt\mathcal{H})\psi^t$ it follows that

$$\frac{d\psi}{dt} = -i\mathcal{H}\psi \tag{9.3}$$

This is the Schrödinger equation. We have allowed ourselves above to use sloppy notations (the "ket" has been omitted). One should realize that the Schrödinger equation reflects the definition of the Hamiltonian as the generator of the evolution.

[9.3] Stationary States (the "Energy Basis")

We can find the eigenstates $|n\rangle$ and the eigenvalues E_n of a Hamiltonian, which is called diagonalization:

$$\mathcal{H}|n\rangle = E_n|n\rangle \tag{9.4}$$

$$\boldsymbol{\mathcal{U}}|n\rangle = \mathrm{e}^{-iE_nt}|n\rangle \tag{9.5}$$

$$\boldsymbol{\mathcal{U}} \rightarrow \delta_{n,m}\mathrm{e}^{-iE_nt} \tag{9.6}$$

Using Dirac notations:

$$\langle n|\boldsymbol{\mathcal{U}}|m\rangle \;=\; \delta_{nm}\mathrm{e}^{-iE_nt} \tag{9.7}$$

If we prepare a superposition of basis states:

$$|\psi^{t=0}\rangle \;=\; \sum_n \psi_n\,|n\rangle \tag{9.8}$$

we get after time t

$$|\psi(t)\rangle \;=\; \sum_n \mathrm{e}^{-iE_nt}\psi_n\,|n\rangle \tag{9.9}$$

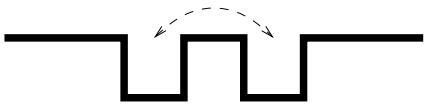
[9.4] How do we construct the Hamiltonian?

We construct the Hamiltonian from "symmetry" considerations. In the next lecture our object will be to show that the Hamiltonian of a non-relativistic particle is of the form:

$$\mathcal{H} \;=\; \frac{1}{2\mathsf{m}}(p-A(x))^2+U(x) \tag{9.10}$$

In this lecture we discuss a simpler case: the Hamiltonian of a particle in a two-site system. We make the following assumptions about the two-site dynamics:

- The system is symmetric with respect to reflection.
- The particle can move from site to site.



These two assumptions determine the form of the Hamiltonian. In addition, we will see how "gauge" considerations can make the Hamiltonian simpler, without loss of generality.

In advance we note that because of gauge considerations, the Hamiltonian can only be determined up to a constant.

$$\mathcal{H} \;\rightarrow\; \mathcal{H}+\epsilon_0\hat{1} \tag{9.11}$$

Namely, if we add a constant to a Hamiltonian, then the evolution operator only changes by a global phase factor:

$$\boldsymbol{\mathcal{U}}(t) \;\rightarrow\; \mathrm{e}^{-it(\mathcal{H}+\epsilon_0\hat{1})} \;=\; \mathrm{e}^{-i\epsilon_0t}\,\mathrm{e}^{-it\mathcal{H}} \tag{9.12}$$

This global phase factor can be gauged away by means of time dependent gauge transformation. We shall discuss gauge transformations in the next sections.

[9.5] The Hamiltonian of a two-site system

It would seem that the most general Hamiltonian for a particle in a two-site system includes 4 parameters:

$$\mathcal{H} = \begin{pmatrix} \epsilon_1 & c\mathrm{e}^{-i\phi} \\ c\mathrm{e}^{i\phi} & \epsilon_2 \end{pmatrix} \tag{9.13}$$

The first observation is that thanks to gauge freedom we can define a new basis such that $\phi = 0$. The new basis is

$$\begin{aligned} |\tilde{1}\rangle &= |1\rangle \\ |\tilde{2}\rangle &= \mathrm{e}^{i\phi}|2\rangle \end{aligned} \tag{9.14}$$

and we see that:

$$\langle\tilde{2}|\mathcal{H}|\tilde{1}\rangle = \mathrm{e}^{-i\phi}\langle 2|\mathcal{H}|1\rangle = \mathrm{e}^{-i\phi}c\mathrm{e}^{i\phi} = c \tag{9.15}$$

Next we can make change the Hamiltonian by a constant ϵ_0. This can be regarded as gauge transformation in time. This means that the basis in time t is identified as $|\tilde{1}\rangle = \exp(-i\epsilon_0 t)|1\rangle$ and $|\tilde{2}\rangle = \exp(-i\epsilon_0 t)|2\rangle$. In this time dependent basis the diagonal matrix elements of the Hamiltonian becomes $\tilde{\epsilon}_1 = \epsilon_1 - \epsilon_0$ and $\tilde{\epsilon}_2 = \epsilon_2 - \epsilon_0$. It should be emphasized that on physical grounds one cannot say whether the old or new basis is "really" time dependent. All we can say is that the new basis is time dependent relative to the old basis. This is just another example of the relativity principle.

The bottom line is that, without loss of generality, we can set $\phi = 0$ and write the most general Hamiltonian of a two site systems with just two physical parameters:

$$\mathcal{H} = \begin{pmatrix} \epsilon & c \\ c & 0 \end{pmatrix} = c\boldsymbol{\sigma}_1 + (\epsilon/2)\boldsymbol{\sigma}_3 + \text{const} = \vec{h}\cdot\vec{\boldsymbol{\sigma}} + \text{const} = h\boldsymbol{\sigma}_n + \text{const} \tag{9.16}$$

where $\vec{h} = (c, 0, \epsilon/2)$, and $h = \sqrt{c^2 + (\epsilon/2)^2}$, such that $\vec{h} = h\vec{n}$, and $\boldsymbol{\sigma}_n = \vec{n}\cdot\vec{\boldsymbol{\sigma}}$. The matrices $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_3$ and in general $\boldsymbol{\sigma}_n$ are known as the Pauli matrices. They satisfy $\boldsymbol{\sigma}_n^2 = 1$, and therefore have eigenvalues ± 1. It follows that the eigenvalues of $\mathcal{H}$ are $\pm h + \text{const}$.

[9.6] The evolution of a two-site system

For the purpose of demonstration we further assume the the two-site system has a reflection symmetry. This implies $\epsilon = 0$. We are left with a single parameter Hamiltonian: In a future lecture we will see that the hopping frequency c is just a different (optional) was to specify the *inertial mass* of the particle. The eigenstates of the mirror symmetric Hamiltonian are the states that are symmetric or anti-symmetric with respect to reflection:

$$|+\rangle = \frac{1}{\sqrt{2}}(|1\rangle + |2\rangle) \tag{9.17}$$

$$|-\rangle = \frac{1}{\sqrt{2}}(|1\rangle - |2\rangle) \tag{9.18}$$

The Hamiltonian in the new basis is:

$$\mathcal{H} = \begin{pmatrix} c & 0 \\ 0 & -c \end{pmatrix} \equiv \begin{pmatrix} E_+ & 0 \\ 0 & E_- \end{pmatrix} \tag{9.19}$$

Let us assume that we have prepared the particle in the first site:

$$|\psi^{t=0}\rangle \;=\; |1\rangle \;=\; \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle) \tag{9.20}$$

The state of the particle, after time t will be:

$$|\psi^{t}\rangle \;=\; \frac{1}{\sqrt{2}}(\mathrm{e}^{-ict}|+\rangle + \mathrm{e}^{-i(-c)t}|-\rangle) \;=\; \cos(ct)|1\rangle - i\sin(ct)|2\rangle \tag{9.21}$$

We see that a particle in a two-site system makes coherent oscillations between the two sites. In particular the probability to find it in the initial site is:

$$P(t) \;=\; |\cos(ct)|^2 \;=\; \frac{1}{2}\left[1+\cos(\Omega t)\right] \tag{9.22}$$

where the frequency of oscillations is $\Omega = 2c$. This result should be contrasted with classical stochastic evolution where the probability to be in each site (if we wait long enough time) would become equal. We can repeat the procedure for a non-symmetric two site system, with the result

$$\Omega \;=\; E_+ - E_- \;=\; \sqrt{(2c)^2+\epsilon^2} \tag{9.23}$$

Here $\pm$ signifies arbitrary indexing of the two eigenstates, namely, $|+\rangle$ is the upper orbital and $|-\rangle$ is the lower orbital.

═══ [9.7] Expectation value

In a later lecture we clarify the notion of *quantum state*. A pure quantum preparation is represented in some sense by a state ψ that provides the information regarding the expected statistics of a measurement: it is the quantum version of *probability function*. We might have an interest in a measurement whose aim is to probe whether that system is in state ϕ. The fundamental postulate of quantum mechanics implies that the probability for that is provided by the *projection formula*:

$$\mathrm{Prob}(\phi|\psi) \;=\; |\langle\phi|\psi\rangle|^2 \tag{9.24}$$

Given an observable A and a state ψ, we define the expectation value $\langle A\rangle$ as follows

$$\langle A\rangle \;=\; \sum_a \mathrm{Prob}(a|\psi)\, a \tag{9.25}$$

where $|a\rangle$ are the eigenstates of A. From the projection formula it follows that

$$\langle A\rangle \;=\; \sum_a |\langle a|\Psi\rangle|^2\, a \tag{9.26}$$

Using linear-algebra identities we get the Sandwich formula:

$$\langle A\rangle \;=\; \langle\psi|A|\psi\rangle \tag{9.27}$$

In a later lecture we clarify that this formula should be regarded as a special case of a more general formula that follows from the basic postulate that defines the notion of *quantum state*.

═══ [9.8] Rate of change of the expectation value

Given an Hamiltonian, with any operator $\hat{A}$ we can associate an operator $\hat{B}$,

$$\hat{B} \;=\; i[\hat{\mathcal{H}}, \hat{A}] \tag{9.28}$$

such that

$$\frac{d\langle \hat{A}\rangle}{dt} \;=\; \langle \hat{B}\rangle \tag{9.29}$$

proof: From the expectation value formula:

$$\langle \hat{A}\rangle_t \;=\; \langle \psi(t)|\hat{A}|\psi(t)\rangle \tag{9.30}$$

Using the Schrödinger equation, we get

$$\begin{aligned} \frac{d}{dt}\Big\langle \hat{A}\Big\rangle &= \Big\langle \frac{d}{dt}\psi\Big|\hat{A}\Big|\psi\Big\rangle + \Big\langle \psi\Big|A\Big|\frac{d}{dt}\psi\Big\rangle \\ &= i\Big\langle \psi\Big|\mathcal{H}A\Big|\psi\Big\rangle - i\Big\langle \psi\Big|A\mathcal{H}\Big|\psi\Big\rangle \;=\; ... \end{aligned} \tag{9.31}$$

═══ [9.9] The semiclassical limit

For the Hamiltonian $H = K(p) + U(x)$ one obtains

$$\frac{d\langle \hat{x}\rangle}{dt} \;=\; \langle i[H,x]\rangle \;=\; \big\langle K'(\hat{p})\big\rangle \;\approx\; K'(\langle \hat{p}\rangle) \tag{9.32}$$

$$\frac{d\langle \hat{p}\rangle}{dt} \;=\; \langle i[H,p]\rangle \;=\; -\big\langle U'(\hat{x})\big\rangle \;\approx\; -U'(\langle \hat{x}\rangle) \tag{9.33}$$

where the approximation hold for a particle whose state is represented by a "minimal" wavepacket, meaning that its position and momentum have very small dispersion. The approximation fails for a particle whose state is extended in phase-space (superposition of distant locations, etc).

═══ [9.10] Beyond the semiclassical limit

Mathematically we can formally regard space (1D for simplicity) as a chain of sites that have separation a. The particle can hop from site to site. In the mathematical limit $a \to 0$ we get the standard Newtonian semiclassical description, where the energy is constant of motion. Consider now a physical potential that divided space into actual sites of size a, as in a metal. Classically the particle (say electron) cannot hop from site to site because there are barriers between the sites. But quantum-mechanically we have to think of the electron as wave. Think of an array of strings in a music instrument. The vibration can move from string to string. This is known as *tunneling*. In a metal the electron can hop from orbital of one atom to the same-energy orbital of the next atom, hence a conduction band is formed.

[10] Dynamics of an N site system

═══ [10.1] Two site system

The simplest non-trivial network that we can imagine is of course a two site system. Without loss of generality we can write the Hamiltonian with a real hopping amplitude c, namely,

$$\mathcal{H} = \begin{pmatrix} \epsilon/2 & c \\ c & -\epsilon/2 \end{pmatrix} \;=\; \frac{\epsilon}{2}\sigma_z + c\sigma_x \;=\; \vec{\Omega}\cdot\vec{S} \tag{10.1}$$

where $\vec{\Omega} = (2c, 0, \epsilon)$. The evolution operator is:

$$\boldsymbol{\mathcal{U}}(t) \;=\; \mathrm{e}^{-it\hat{\mathcal{H}}} \;=\; \mathrm{e}^{-i(\vec{\Omega}t)\cdot\vec{S}} \;=\; R(\vec{\Phi}(t)) \tag{10.2}$$

where $\vec{\Phi}(t) = \vec{\Omega}t$. This means that the spin makes precession. It is best to represent the state of the spin using the polarization vector $\vec{M} = (\langle\sigma_x\rangle, \langle\sigma_y\rangle, \langle\sigma_z\rangle)$. Then we can describe the precession using a classical-like "Bloch sphere picture". Considering a symmetric two site system ($\epsilon = 0$) we can find the eigenstates, and then find the evolution by expanding the initial state at that basis. The frequency of the oscillations equals to the energy splitting of the eigen-energies. But once ($\epsilon \neq 0$) this scheme becomes very lengthy and intimidating. It is much easier to use the analogy with spin 1/2. Then it is clear, just by looking at the Hamiltonian that the oscillation frequency is

$$\Omega \;=\; \sqrt{(2c)^2 + \epsilon^2} \tag{10.3}$$

and hence the eigenenergies are $E_\pm = \pm\Omega/2$. Furthermore, it is clear that the precession axis is tilted relative to the z axis with an angle

$$\theta \;=\; \arctan(2c/\epsilon) \tag{10.4}$$

Accordingly the eigenstates can be obtained from the $\uparrow$ and from the $\downarrow$ states by rotating them an angle θ around the y axis:

$$|+\rangle = \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix}, \qquad\qquad |-\rangle = \begin{pmatrix} -\sin(\theta/2) \\ \cos(\theta/2) \end{pmatrix} \tag{10.5}$$

If we are interested in studying the dynamics there is no need to expand the initial state in this basis. It is much easier to write the explicit expression for the evolution operator:

$$\boldsymbol{\mathcal{U}}(t) \;=\; R(\vec{\Phi}(t)) \;=\; \cos(\Omega t/2) - i\sin(\Omega t/2)\Big[\cos\theta\ \boldsymbol{\sigma}_z + \sin\theta\ \boldsymbol{\sigma}_x\Big] \tag{10.6}$$

Let us assume that initially the system is in the "up" state. We get that the state after time t is

$$|\psi(t)\rangle \;=\; \boldsymbol{\mathcal{U}}(t)|\uparrow\rangle \;=\; \Big[\cos(\Omega t/2) - i\cos\theta\sin(\Omega t/2)\Big]|\uparrow\rangle - i\Big[\sin\theta\sin(\Omega t/2)\Big]|\downarrow\rangle \tag{10.7}$$

Let us define $P(t)$ as the probability to be found after time t in the same state. Initially we have $P(0) = 1$. We can easily find the explicit expression for $P(t)$ without having to do any diagonalization of the Hamiltonian. Rather we exploit the explicit expression above and get

$$P(t) \quad = \quad \left| \langle \uparrow | \psi(t) \rangle \right|^2 \quad = \quad 1 - \sin^2(\theta) \sin^2 \left(\frac{\Omega t}{2} \right) \tag{10.8}$$

This result is called the Rabi Formula. We see that in order to have nearly complete transitions to the other site after half a period, we need a very strong coupling ($c \gg \epsilon$). In the opposite limit ($c \ll \epsilon$) the particle tends to stay in the same site, indicating that the eigenstates are barely affected.

[10.2] N site chain

Of particular important is the the standard tight binding model where we have a chain of N sites with near neighbor hopping. For simplicity we assume periodic boundary conditions, which is like saying that we have a closed ring. The sites are positioned at $x = a \times$ integer. The Hamiltonian, with real hopping amplitudes $-c$, can be written as

$$\mathcal{H} \quad = \quad -c(\hat{D} + \hat{D}^{-1}) + U(\hat{x}) \quad = \quad -2c \cos(a\hat{p}) + U(\hat{x}) \tag{10.9}$$

where D is the one site displacement operator. The definition of the velocity operator is implied by the "rate of change formula", namely

$$\hat{v} \quad = \quad i[\mathcal{H}, \hat{x}] \quad = \quad ica(\hat{D} - \hat{D}^{-1}) \quad = \quad 2ca \sin(a\hat{p}) \tag{10.10}$$

Note that for small velocities we have linear dispersion relation $v = (1/\mathsf{m})p$, where the mass is $\mathsf{m} = 1/(2ca^2)$. The small a continuum limit of the above Hamiltonian is the standard non-relativistic Hamiltonian. For a full presentation of the standard non-relativistic Hamiltonian is 3D, including magnetic field, see lectures 9 and 10 of arXiv:quant-ph/0605180.

[10.3] Applying a force on a particle

Consider particle whose coordinate x describes its position in the lattice or along an axis. Let us add to the Hamiltonian a term $U(x) = -fx$. The meaning of such term is that the particle experiences a uniform field of force. The classical equation of motion $\dot{p} = f$ implies that the momentum increases, namely $p = p_0 + ft$. Let us verify that the same result strictly holds also quantum mechanically:

$$e^{-iU(x)t} |p_0\rangle \quad = \quad e^{ift\hat{x}} \sum_x e^{ip_0 x} |x\rangle \quad = \quad \sum_x e^{i(p_0+ft)x} |x\rangle \quad = \quad |p_0 + ft\rangle \tag{10.11}$$

So both quantum mechanics and classical mechanics agree that the momentum is growing. What about the location of the particle? Here it is important to distinguish between motion along chain and motion along axis. The dispersion relation is not the same. The equation of motion is $\dot{x} = 2ca \sin(ap)$ in the lattice case, as opposed to $\dot{x} = (1/\mathsf{m})p$ in the continuum case. Therefore, instead of getting accelerated motion in space, we get so called *Bloch oscillations*, which can be regarded as generalization of two-site *Rabi oscillations*.

Note that the room-temperature motion of electrons in metals is very different. We have neither acceleration nor oscillations. Rather we have diffusion and drift, with drift velocity that is proportional to the force. This reflects stochastic dynamics due to the surrounding environment.

[11] Theory of quantum computation

═══ [11.1] Motivating Quantum Computation

Present day secure communication is based on the RSA two key encryption method. The RSA method is based on the following observation: Let N be the product of two unknown big prime numbers p and q. Say that we want to find are what are its prime factors. The simple minded way would be to try to divide N by 2, by 3, by 5, by 7, and so on. This requires a huge number ($\sim N$) of operations. It is assumed that N is so large such that in practice the simple minded approached is doomed. Nowadays we have the technology to build classical computers that can handle the factoring of numbers as large as $N \sim 2^{300}$ in a reasonable time. But there is no chance to factor larger numbers, say of the order $N \sim 2^{308}$. Such large numbers are used by Banks for the encryption of important transactions. In the following sections we shall see that factoring of a large number N would become possible once we have a quantum computer.

Computational complexity: A given a number N can be stored in an n-bit register. The size of the register should be $n \sim \log(N)$, rounded upwards such that $N \leq 2^n$. As explained above in order to factor a number which is stored in an n bit register by a classical computer we need an exponentially large number ($\sim N$) of operations. Obviously we can do some of the operations in parallel, but then we need an exponentially large hardware. Our mission is to find an efficient way to factor an n-bit number that do not require exponentially large resources. It turns out that a quantum computer can do the job with hardware/time resources that scale like power of n, rather than exponential in n. This is done by finding a number N_2 that has a common divisor with N. Then it is possible to use Euclid's algorithm in order to find this common divisor, which is either p or q.

Euclid's algorithm: There is no efficient algorithm to factor a large number $N \sim 2^n$. The classical computational complexity of this task is exponentially large in n. But if we have two numbers $N_1 = N$ and N_2 we can quite easily and efficiently find their greater common divisor $\mathrm{GCD}(N_1, N_2)$ using Euclid's algorithm. Without loss of generality we assume that $N_1 > N_2$. The two numbers are said to be co-prime if $\mathrm{GCD}(N_1, N_2) = 1$. Euclid's algorithm is based on the observation that we can divide N_1 by N_2 and take the reminder $N_3 = \mathrm{mod}(N_1, N_2)$ which is smaller than N_2. Then we have $\mathrm{GCD}(N_1, N_2) = \mathrm{GCD}(N_2, N_3)$. We iterate this procedure generating a sequence $N_1 > N_2 > N_3 > N_4 > \cdots$ until the reminder is zero. The last non-trivial number in this sequence is the greater common divisor.

The RSA encryption method: The RSA method is based on the following mathematical observation. Given two prime numbers p and q define $N = pq$. Define also a and b such that $ab = 1 \mod [(p-1)(q-1)]$. Then we have the relations

$$B = A^a \mod [N] \tag{11.1}$$

$$A = B^b \mod [N] \tag{11.2}$$

This mathematical observation can be exploited as follows. Define

$$\text{public key} = (N, a) \tag{11.3}$$

$$\text{private key} = (N, b) \tag{11.4}$$

If anyone want to encrypt a message A, one can use for this purpose the public key. The coded message B cannot be decoded unless one knows the private key. This is based on the assumption that the prime factors p and q and hence b are not known.

[11.2] The factoring algorithm

According to Fermat's theorem, if N is prime, then $M^N = M \mod (N)$ for any number $M(< N)$. If the "seed" M is not divisible by N, this can be re-phrased as saying that the period of the function $f(x) = M^x \mod (N)$ is $r = N-1$. That means $f(x+r) = f(x)$. To be more precise the primitive period can be smaller than N (not any "seed" is "generator"). More generally if N is not prime, and the seed M has no common divisor with it, then the primitive period of $f(x)$ is called "the order".

The quantum computer is a black box that allows to find the period r of a function $f(x)$. How this is done will be discussed in the next section. Below we explain how this ability allows us to factor a given large number N.

(1) We have to store N inside an n-bit register.

(2) We pick a large number M, so called seed, which is smaller than N. We assume that M is co-prime to N. This assumption can be easily checked using Euclid'd algorithm. If by chance the chosen M is not co-prime to N then we are lucky and we can factor N without quantum computer. So we assume that we are not lucky, and M is co-prime to N.

(3) We build a processor that can calculate the function $f(x) = M^x \mod (N)$. This function has a period r which is smaller than N.

(4) Using a quantum computer we find one of the Fourier components of $f(x)$ and hence its period r. This means that $M^r = 1 \mod (N)$.

(5) If r is not even we have to run the quantum computer a second time with a different M. Likewise if $a^{r/2} = -1$. There is a mathematical theorem that guarantees that with probability of order one we should be able to find M with which we can continue to the next step.

(6) We define $Q = M^{r/2} \mod (N)$. We have $Q^2 = 1 \mod (N)$, and therefore $(Q-1)(Q+1) = 0 \mod (N)$. Consequently both $(Q-1)$ and $(Q+1)$ must have either p or q as common divisors with N.

(6) Using Euclid's algorithm we find the GCD of N and $\tilde{N} = (Q-1)$, hence getting either p or q.

The bottom line is that given N and M an an input, we would like to find the period r of the functions

$$f(x) = M^x \mod (N) \tag{11.5}$$

Why do we need a quantum computer to find the period? Recall that the period is expected to be of order N. Therefore the x register should be n_c bits, where n_c is larger or equal to n. Then we have to make order of 2^{n_c} operations for the purpose of evaluating $f(x)$ so as to find out its period. It is assumed that n is large enough such that this procedure is not practical. We can of course try to do parallel computation of $f(x)$. But for that we need hardware which is larger by factor of 2^n. It is assumed that to have such computational facility is equally not practical. We say that factoring a large number has an exponentially complexity.

The idea of quantum processing is that the calculation of $f(x)$ can be done "in parallel" without having to duplicate the hardware. The miracle happens due to the superposition principle. A single quantum register of size n_c can be prepared at $t = 0$ with all the possible input x values in superposition. The calculation of $f(x)$ is done in parallel on the prepared state. The period of $f(x)$ in found via a Fourier analysis. In order to get good resolution n_c should be larger than n so as to have $2^{n_c} \gg 2^n$. Neither the memory, nor the number of operation, are required to be exponentially large.

[11.3] The quantum computation architecture

We shall regard the computer as a machine that has registers for memory and gates for performing operations. The complexity of the computation is determined by the total number of memory bits which has to be used times the number of operations. In other words the complexity of the calculation is the product memory × time. As discussed in the previous section classical parallel computation do not reduce the complexity of a calculation. Both classical and quantum computers work according to the following scheme:

$$|\text{output}\rangle \;=\; U[\text{input}]\;|0\rangle \tag{11.6}$$

This means that initially we set all the bits or all the qubits in a zero state (a reset operation), and then we operate on the registers using gates, taking the input into account. It is assumed that there is a well define set of elementary gates. An elementary gate (such as "AND") operates on few bits (2 bits in the case of AND) at a time. The size of the hardware is determined by how many elementary gates are required in order to realize the whole calculation.

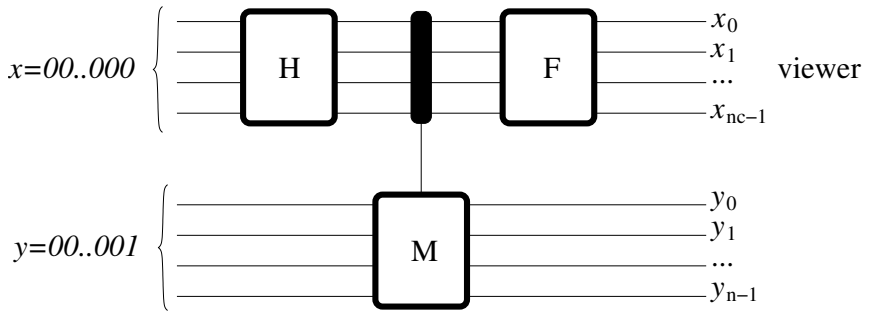
The quantum analog of the digital bits ("0" or "1") are the qubits, which can be regarded as spin 1/2 elements. These are "analog" entities because the "spin" can point in any direction (not just "up" or "down"). The set of states such that each spin is aligned either "up" or "down" is called the *computational basis*. Any other state of a register can be written a superposition:

$$|\Psi\rangle \;=\; \sum_{x_0,x_1,x_2,...} \psi(x_0,x_1,x_2...)\;|x_0,x_1,x_2,...\rangle \tag{11.7}$$

The architecture of the quantum computer which is requited in order to find the period r of the function $f(x)$ is illustrated in the figure below. We have two registers:

$$x \;=\; (x_0,x_1,x_2,...,x_{n_c-1}) \tag{11.8}$$
$$y \;=\; (y_0,y_0,y_2,...,y_{n-1}) \tag{11.9}$$

The registers x and y can hold binary numbers in the range $x < N_c$ and $y < \bar{N}$ respectively, where $N_c = 2^{n_c}$ and $\bar{N} = 2^n > N$. The y register is used by the CPU for processing mod(N) operations and therefore it requires a minimal number of n bits. The x register has n_c bits and it is used to store the inputs for the function $f(x)$. In order to find the period of $f(x)$ the size n_c of the latter register should be significantly larger compared with n. Note that that $n_c = n + 10$ implies that the x range becomes roughly ×1000 larger than the expected period. Large n_c is required if we want to determine the period with large accuracy.



We are now ready to describe the quantum computation. In later sections we shall give more details, and in particular we shall see that the realization of the various unitary operations which are listed below does not

require an exponentially large hardware. The preliminary stage is to make a "reset" of the registers, so as to have

$$|\Psi\rangle = |x;y\rangle = |0,0,0,...,0,0;1,0,0,...,0,0\rangle \tag{11.10}$$

Note that it is convenient to start with $y = 1$ rather than $y = 0$. Then come a sequence of unitary operations

$$U = U_F U_M U_H \tag{11.11}$$

where

$$U_H = U_{\text{Hadamard}} \otimes \mathbf{1} \tag{11.12}$$
$$U_M = \sum_x |x\rangle\langle x| \otimes U_M^{(x)} \tag{11.13}$$
$$U_F = U_{\text{Fourier}} \otimes \mathbf{1} \tag{11.14}$$

The first stage is a unitary operation U_H that sets the x register in a democratic state. It can realized by operating on Ψ with the Hadamard gate. Disregarding normalization we get

$$|\Psi\rangle = \sum_x |x\rangle \otimes |y{=}1\rangle \tag{11.15}$$

The second stage is an x controlled operation U_M. This stage is formally like a quantum measurement: The x register is "measured" by the y register. The result is

$$|\Psi\rangle = \sum_x |x\rangle \otimes |y{=}f(x)\rangle \tag{11.16}$$

Now the y register is entangled with the x register. The fourth stage is to perform Fourier transform on the x register:

$$|\Psi\rangle = \sum_x \left[\sum_{x'} \mathrm{e}^{i\frac{2\pi}{N_c}xx'}|x'\rangle\right] \otimes |f(x)\rangle \tag{11.17}$$

We replace the dummy integer index x' by $k = (2\pi/N_c)x'$ and re-write the result as

$$|\Psi\rangle = \sum_k |k\rangle \otimes \left[\sum_x \mathrm{e}^{ikx}|f(x)\rangle\right] \equiv \sum_k p_k|k\rangle \otimes |\chi^{(k)}\rangle \tag{11.18}$$

The final stage is to measure the x register. The probability to get k as the result is

$$p_k = \left|\left|\sum_x \mathrm{e}^{ikx}|f(x)\rangle\right|\right|^2 \tag{11.19}$$

The only non-zero probabilities are associated with $k = \text{integer} \times (2\pi/r)$. Thus we are likely to find one of these k values, from which we can deduce the period r. Ideally the error is associated with the finite length of the x register. By making the x register larger the visibility of the Fourier components becomes better.

[11.4] Elementary qubit quantum gates

Classical operations can be constructed using single bit NOT gates, and two-bit AND gates. Similarly, quantum operations can be constructed using single qubit operations (rotations), and two-qubit CNOT gates. We review below the common gates that are used, e.g. in the IBM composer.

The single qubit gates can be regarded as spin rotations. It is useful to define the following notations for the observable that 'measures' the state of the qubit in the standard basis:

$$Q \;=\; \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \;=\; \frac{1}{2}(1-Z) \;=\; |1\rangle\langle 1| \tag{11.20}$$

The notations X, Y, Z are used for the Pauli matrices. The notations RX, RY, RZ are used for rotations. A phase gate is defined as follows:

$$\mathrm{P}(\varphi) \;=\; \mathrm{e}^{i\varphi Q} \;=\; \mathrm{e}^{i(\varphi/2)\mathsf{RZ}(\varphi)} \;=\; |0\rangle\langle 0| + |1\rangle\mathrm{e}^{i\varphi}\langle 1| \tag{11.21}$$

Note that

$$Z \;=\; \mathrm{P}(\pi) \;=\; i\mathrm{RZ}(\pi) \tag{11.22}$$

More interesting are two-qubit gates. The controlled-phase gate is

$$\mathrm{CP}(\varphi) \;=\; \mathrm{e}^{i\varphi Q\otimes Q} \;=\; |0\rangle\langle 0|\otimes 1 + |1\rangle\langle 1|\otimes \mathrm{P}(\varphi) \tag{11.23}$$

This gate is symmetric: we can say that the first qubit controls the phase operation on the second qubit, or vice versa. In particular the controlled-Z gate is

$$\mathrm{CZ} \;=\; \mathrm{CP}(\pi) \;=\; \mathrm{e}^{i\pi Q\otimes Q} \;=\; |0\rangle\langle 0|\otimes 1 + |1\rangle\langle 1|\otimes Z \tag{11.24}$$

Most useful is the controlled-X (CNOT) gate

$$\mathrm{CX} \;=\; |0\rangle\langle 0|\otimes 1 + |1\rangle\langle 1|\otimes \mathrm{X} \;=\; \begin{pmatrix} 1 & 0 & & \\ 0 & 1 & & \\ & & 0 & 1 \\ & & 1 & 0 \end{pmatrix} \tag{11.25}$$

Ising-type phase gates are defined as follows:

$$\mathrm{RZZ}(\varphi) \;=\; \mathrm{e}^{-i(\varphi/2)Z\otimes Z} \tag{11.26}$$

Similarly one can define RXX and RYY gates, etc.

For completeness we list some common notations:

$$\begin{aligned}
T &= P(\pi/4) = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{e}^{i\pi/4} \end{pmatrix} \\
S &= T^2 = P(\pi/2) = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \\
Z &= S^2 = P(\pi) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \sigma_z = i\mathrm{RZ}(\pi) \\
X &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \sigma_x = i\mathrm{e}^{-i\pi S_x} = i\mathrm{RX}(\pi) = \text{NOT gate} \\
Y &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \sigma_y = i\mathrm{e}^{-i\pi S_y} = i\mathrm{RY}(\pi) \\
H &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{\sqrt{2}}(\sigma_x + \sigma_z) = i\mathrm{e}^{-i\pi S_n} = \text{Hadamard gate} \\
R &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \frac{1}{\sqrt{2}}(1 - i\sigma_y) = \mathrm{e}^{-i(\pi/2)S_y} = \mathrm{RY}(\pi/2) = 90\text{deg Rotation}
\end{aligned} \tag{11.27}$$

We have $R^4 = -1$ which is a 2π rotation in $SU(2)$. We have $X^2 = Y^2 = Z^2 = H^2 = 1$ which implies that these are π rotations in $U(2)$. We emphasize that though the operation of H on "up" or "down" states looks like $\pi/2$ rotation, it is in fact a π rotation around a 45^o inclined axis:

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right) \cdot \vec{\sigma} = \vec{n} \cdot \vec{\sigma} \tag{11.28}$$

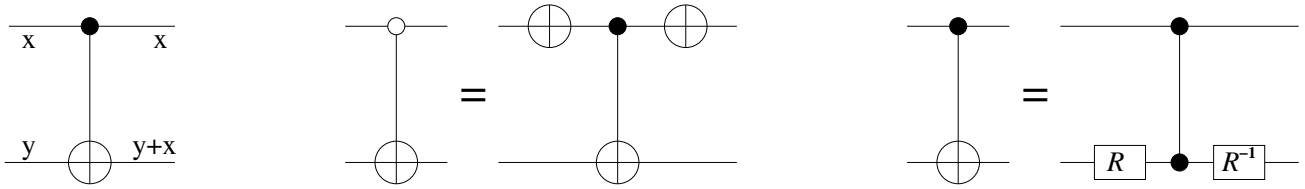
Global phase issue.– Mathematically we can define a U(1) operation on a qubit, namely, a multiplication by a phase. Let us use the notation $\mathrm{G}(\varphi) = e^{i\varphi}$. Clearly, from a physical perspective this operation has no significance. Operations like X and $\mathrm{RX}(\pi)$ are equivalent. If we have, say, a two-qubit register, then $[G(\varphi)\,|x_0\rangle] \otimes |x_1\rangle = e^{i\varphi}\,|x_0, x_1\rangle$. What about a controlled-G? For that we have the identity

$$\mathrm{CG}(\varphi) \;=\; |0\rangle\langle 0| \otimes 1 + |1\rangle\langle 1| \otimes e^{i\varphi} \;=\; \mathrm{P}(\varphi) \otimes 1 \tag{11.29}$$

Namely, CG is mathematically equivalent to an operation with a phase-gate on the control qubit. For that reason the G-gate is redundant.

[11.5] The CNOT quantum gate

It is popular to consider the controlled NOT (CNOT) gate as an elementary gate, with which we can construct more complicated gates. We shall use for it the notation CX. In the standard computational basis the control bit (x) of CX is not affected by its operation, while the controlled bit (y) undergoes NOT provided the x bit is turned-*on*. The gate is schematically illustrated in the following figure. We also provide in the figure another variation of CX that performs NOT provided the x bit is turned-*off*. Below we explain that CX can be constructed using CZ and R operations (right panel of the figure).



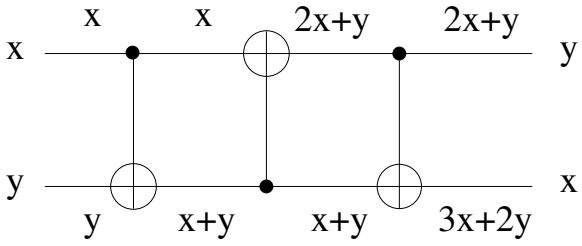
In order to realize a two qubit gate we need a coupling mechanism. If we regard the two qubits as spins, it is most natural to exploit the Ising spin-spin interaction ($\propto Z \otimes Z$) in order to realize an RZZ phase operation. But for pedagogical purpose it is more convenient to regard the qubits as dots that can be charges with $Q = 0, 1$ particles, such that the interaction is $\propto Q \otimes Q$. With such interaction we can construct a CZ operation. Then we can exploit the similarity transformation $R^{-1} Z R = X$ in order to get CX as follows:

$$\text{CX} \;=\; R^{-1}\,\text{CZ}\,R \;=\; |0\rangle\langle 0| \otimes 1 + |1\rangle\langle 1| \otimes X \tag{11.30}$$

where the rotations are performed on the controlled qubit (see figure). Here R is -90deg rotation around Y.

[11.6] CNOT-related operations

Having the ability to realize single-qubit operations, and the CNOT gate, we can construct all other possible gates and circuits. It is amusing to see how SWAP gate can be realized by combining 3 CNOT gates, as illustrated is the following diagram:



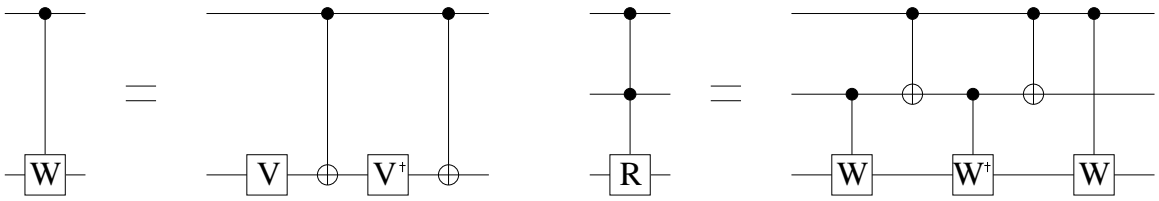
The controlled-not gate has misc generalizations. There are two aspects: (i) how to form controlled-rotations instead of control-not; (ii) how to have more then one control bit. The way to form a controlled RY or controlled RZ rotation is to write the rotation, call it W, as a sequence of two half-rotations, namely, $W = VV$. We use the notation $\bar{V} = V^\dagger = V^{-1}$ for the inverse rotation. We note that for either RY or RZ we have the identity $XVX = \bar{V}$. Then we can use the following circuit (see figure below for its illustration):

$$\text{CW} \;=\; (\text{CX}\;\bar{V}\;\text{CX})\;V \tag{11.31}$$

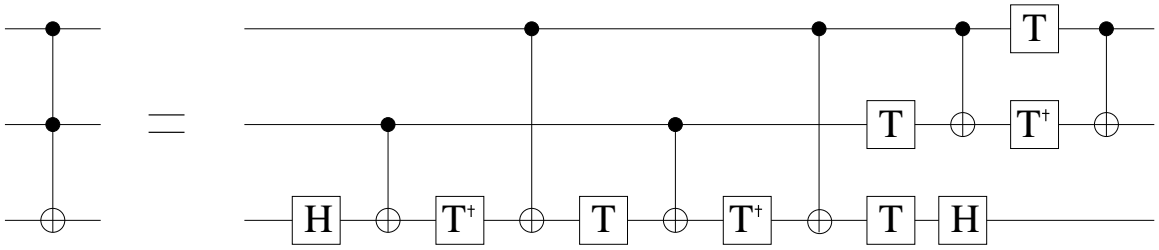
where V operates on the controlled qubit, and the CX performs on it a controlled-not. Note that if we want to get CRX, we can use the identity CRX = RZ$(-\pi/2)$ CRY RZ$(\pi/2)$, where the RZ rotations operate on the controlled qubit. The way to have two control bits labelled as "1" and "2" is to write the desired rotation as $R = WW$, and then to use the following circuit (see figure below for its illustration):

$$\text{CCR} \quad = \quad \text{CW}^{2\mapsto 0}\ \text{CX}^{2\mapsto 1}\ \text{C}\bar{\text{W}}^{1\mapsto 0}\ \text{CX}^{2\mapsto 1}\ \text{CW}^{1\mapsto 0} \tag{11.32}$$

where the notation $\text{C}^{i\mapsto j}$ means that qubit i controls the operation on qubit j.



The CCX operation, where we have two qubit that control NOT operation, is known as Toffoli gate. The NOT operation is performed only if both control bits are turned on:



The realization of the Toffoli gate opens the way to the quantum realization of an AND operation. Namely, by setting $y = 0$ at the input, the output would be $y = x_1 \wedge x_2$. For generalization of the Toffoli gate to the case where the x register is of arbitrary size see p.184 of Nielsen and Chuang.

[11.7] Quantum Simulations

We have motivated quantum computation by considering the RSA decryption problem. We shall come back to that in the subsequent sections, providing details on the Hadamard transform, on the quantum Fourier transform, and on the required controlled computation of the $f(x)$ function.

But quantum computation can be used also for the purpose of quantum simulations. Clearly, if we have a large N-spin system, it is not feasible to perform a classical simulation with $2^N \times 2^N$ matrices, while it might be feasible to perform a simulation with an N qubit register. If we consider spin 1/2 system, then each qubit can represent a spin. Evolution that is generated by, say, $S_x \otimes S_x$ interaction can be realized using RXX gates. The Hamiltonian might be a sum of several such terms, and therefore the use of the Trotter formula is required: each dt time step of the simulation is composed of a sequence of infinitesimal unitary operations that correspond to the terms in the Hamiltonian.

Slightly less natural is to perform simulation of a 2^N site system, where each site is labelled by an N digit binary number $\boldsymbol{x}$. The operation of a potential $V(\boldsymbol{x})$ can be simulated using controlled-phase operations. Note: if we apply a phase gate on a single qubit, say x_0, then all the sites that have $x_0 = 1$ are affected. If we want to

change the phase of, say, $\boldsymbol{x} = \cdots 001$, we need a controlled operation that performs the phase operation only if all the other bits are 0. Similar treatment is required in order to induce hopping. Flipping one bit in $\boldsymbol{x}$, say x_0, means simultaneous transitions within any pair of sites whose $\boldsymbol{x}$ differ by digit x_0. If we want to induce transition between, say, only the first two sites ($\cdots 000$ and $\cdots 001$) we have to use a controlled operation on the x_0 bit, namely, it is "rotated" only if all the other bits equal 0. Accordingly, with controlled operation we can induce any transition between two sites whose $\boldsymbol{x}$ differs by a single binary digit. In order to induce transitions between sites whose $\boldsymbol{x}$ differ by more than one digit, we have to perform Gray-sequence of exchanges before the controlled operation, and after that reverse this sequence. For Hamiltonian of the type $H = T(\boldsymbol{p}) + V(\boldsymbol{x})$ we might use the identity $T(\boldsymbol{p}) = F^{\dagger} T(\boldsymbol{x}) F$, where F is the quantum Fourier transform operation (see below). Other trick can be used for e.g. simulating the effect of $T(\boldsymbol{p}) = \cos(\boldsymbol{p})$.

═══ [11.8] The Hadamard Transform

In the following we discuss the Hadamard and the Fourier transforms. These are unitary operations that are defined on the multi-qubit x register. A given basis state $|x_0, x_1, x_3, ...\rangle$ can be regarded as the binary representation of an integer number:

$$x = \sum_{r=0}^{n_c - 1} x_r 2^r \tag{11.33}$$

We distinguish between the algebraic multiplication for which we use the notation xx', and the scalar product for which we use the notation $x \cdot x'$,

$$\begin{aligned} x \cdot x' &= \sum_r x_r x'_r \\ xx' &= \sum_{r,s} x_r x'_s 2^{r+s} \end{aligned} \tag{11.34}$$

So far we have defined the single qubit Hadamard gate. If we have an multi-qubit register it is natural to define

$$U_{\text{Hadamard}} = H \otimes H \otimes H \otimes \cdots \tag{11.35}$$

The operation of a single-qubit Hadamard gate can be written as

$$|x_1\rangle \xrightarrow{H} \frac{1}{\sqrt{2}} \left(|0\rangle + (-1)^{x_1} |1\rangle \right) = \frac{1}{\sqrt{2}} \sum_{k_1 = 0,1} (-1)^{k_1 x_1} |k_1\rangle \tag{11.36}$$

If we have a multi-qubit register we simply have to perform (in parallel) an elementary Hadamard transform on each qubit:

$$\begin{aligned} |x_0, x_1, ..., x_r, ...\rangle \xrightarrow{H} \quad \prod_r \frac{1}{\sqrt{2}} \left(|0\rangle + (-1)^{x_r} |1\rangle \right) &= \frac{1}{\sqrt{2^{n_c}}} \prod_r \left[\sum_{k_r = 0,1} (-1)^{k_r x_r} |k_r\rangle \right] \\ = \frac{1}{\sqrt{N_c}} \sum_{k_0, k_1, ...} (-1)^{k_0 x_0 + k_1 x_1 + \cdots} |k_0, k_1, ..., k_r, ...\rangle &= \frac{1}{\sqrt{N_c}} \sum_k (-1)^{k \cdot x} |k\rangle \end{aligned} \tag{11.37}$$

The Hadmard transform is useful in order to prepare a "democratic" superposition state as follows:

$$|0,0,...,0\rangle \quad \stackrel{H}{\longrightarrow} \quad \frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)\otimes\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)\otimes...\otimes\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle) \quad \mapsto \quad \frac{1}{\sqrt{N_c}}\begin{pmatrix}1\\1\\.\\.\\.\\1\end{pmatrix} \tag{11.38}$$

To operate with a unitary operator on this state is like making parallel computation on all the possible x basis states.

[11.9] The quantum Fourier transform

The definitions of the Hadamard transform and the quantum Fourier transform are very similar in style:

$$U_{\text{Hadamard}}|x\rangle = \frac{1}{\sqrt{N_c}}\sum_k(-1)^{k\cdot x}|k\rangle \tag{11.39}$$

$$U_{\text{Fourier}}|x\rangle = \frac{1}{\sqrt{N_c}}\sum_k \mathrm{e}^{-i\frac{2\pi}{N_c}kx}|k\rangle \tag{11.40}$$

Let us write the definition of the quantum Fourier transform using different style so as to see that it is indeed a Fourier transform operation in the conventional sense. First we notice that its matrix representation is

$$\langle x'|U_{\text{Fourier}}|x\rangle = \frac{1}{\sqrt{N_c}}\mathrm{e}^{-i\frac{2\pi}{N_c}x'x} \tag{11.41}$$

If we operate with it on the state $|\psi\rangle=\sum_x\psi_x|x\rangle$ we get $|\varphi\rangle=\sum_x\varphi_x|x\rangle$, where the column vector φ_x is obtained from ψ_x by a multiplication with the matrix that represents U_{Fourier}. Changing the name of the dummy index form x to k we get the relation

$$\varphi_k = \frac{1}{\sqrt{N_c}}\sum_{x=0}^{N_c-1}\mathrm{e}^{-i\frac{2\pi}{N_c}kx}\psi_x \tag{11.42}$$

This is indeed the conventional definition of

$$\begin{pmatrix}\psi_0\\\psi_1\\\psi_2\\.\\.\\.\\\psi_{N_c-1}\end{pmatrix} \quad \stackrel{FT}{\longrightarrow} \quad \begin{pmatrix}\varphi_0\\\varphi_1\\\varphi_2\\.\\.\\.\\\varphi_{N_c-1}\end{pmatrix} \tag{11.43}$$

The number of memory bits which are required to store these vectors in a classical register is of order $N\sim 2^n$. The number of operations which is involved in the calculation of a Fourier transform seems to be of order N^2. In fact

there is an efficient "Fast Fourier Transform" (FFT) algorithm that reduces the number of required operations to $N \log N = n2^n$. But this is still an exponentially large number in n. In contrast to that a quantum computer can store these vectors in n qubit register. Furthermore, the "quantum" FT algorithm can perform the calculation with only $n^2 \log n \log\log n$ operations. We shall not review here how the Quantum Fourier transform is realized. This can be found in the textbooks. As far as this presentation is concerned the Fourier transform can be regarded as a complicated variation of the Hadamard transform.

═══ [11.10] Note on analog or optical computation

A few words are in order here regarding quantum computation versus classical analog computation. In an analog computer every analog "bit" can have a voltage within some range, so ideally each analog bit can store infinite amount of information. This is of course not the case in practice because the noise in the circuit defines some effective finite resolution. Consequently the performance is not better compared with a digital computers. In this context the analog resolution is a determining factor in the definition of the memory size. Closely related is optical computation. This can be regarded as a special type of analog computation. Optical Fourier Transform of a "mask" can be obtained on a "screen" that is placed in the focal plane of a lens. The FT is done in one shot. However, also here we have the same issue: Each pixel of the mask and each pixel of the screen is a hardware element. Therefore we still need an exponentially large hardware just to store the vectors. At best the complexity of FT with optical computer is of order 2^n.

═══ [11.11] The U_M operation

The CNOT/Toffoli architecture can be generalized so as to realize any operation of the type $y = f(x_1, x_2, ...)$, as an x-controlled operation, where y is a single qubit. More generally we have

$$x = (x_0, x_1, x_2 ... x_{n_c-1}) \tag{11.44}$$
$$y = (y_0, y_1, y_2, ... y_{n-1}) \tag{11.45}$$

and we would like to realize a unitary controlled operation

$$U = \sum_x |x\rangle\langle x| \otimes U^{(x)} \equiv P^{(0)} \otimes U^{(0)} + P^{(1)} \otimes U^{(1)} + P^{(2)} \otimes U^{(2)} + ... \tag{11.46}$$

This is formally like a measurement of the x register by the y register. Note that x is a constant of motion, and that U has a block diagonal form:

$$\langle x', y'|U|x, y\rangle = \delta_{x',x} U^{(x)}_{y',y} = \begin{pmatrix} U^{(0)} & & & \\ & U^{(1)} & & \\ & & U^{(2)} & \\ & & & ... \end{pmatrix} \tag{11.47}$$

Of particular interest is the realization of a unitray operation that maps $y = 1$ to $y = f(x)$. Let us look on

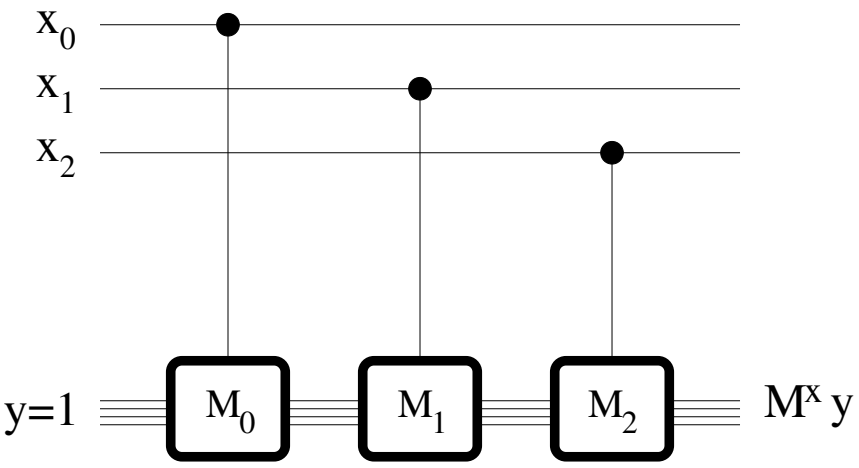
$$U^{(x)}_M \Big| y \Big\rangle = \Big| M^x y \mod (N) \Big\rangle \tag{11.48}$$

If M is co-prime to N, then U is merely a permutation matrix, and therefore it is unitary. The way to realize this

operation is implied by the formula

$$M^x \;=\; M^{\sum_s x_s 2^s} \;=\; \prod_s \left(M^{2^s}\right)^{x_s} \;=\; \prod_{s=0}^{n_c-1} M_s^{x_s} \tag{11.49}$$

which requires n_c stages of processing. The circuit is illustrated in the figure below. In the s stage we have to perform a controlled multiplication of y by $M_s \equiv M^{2^s} \mod (N)$.



Fundamentals

[12] The uncertainty principle

Light version of this lecture in Hebrew is availble in the course site. Full version of this lecture is available in arXiv:quant-ph/0605180, which includes discussion regarding the notion of quantum state, the basic postulates of the theory, and the representation using a probability matrix.

[12.1] The two slit experiment

If we have a beam of electrons, that have been prepared with a well defined velocity, and we direct it to a screen through two slits, then we get an interference pattern from which we can determine the "de-Broglie wavelength" of the electrons. I will assume that the student is familiar with the discussion of this experiment from introductory courses. The bottom line is that the *individual* electrons behave *like wave* and can be characterized by a wavefunction $\psi(x)$. This by itself does not mean that our world in not classical. We still can speculate that $\psi(x)$ has a classical interpenetration. Maybe our modeling of the system is not detailed enough. Maybe the two slits, if they are both open, deform the space in a special way that makes the electrons likely to move only in specific directions? Maybe, if we had better experimental control, we could *predict with certainty* where each electron will hit the screen.

The modern interpenetration of the two slit experiment is not classical. The so called "quantum picture" is that the electron can be at the same time at two different places: it goes via both slits and interferes with itself. This sounds strange. Whether the quantum interpenetration is correct we cannot establish: maybe in the future we will have a different theory. What we *can* establish is that a classical interpretation of reality is *not* possible. This statement is based on a different type of an experiment that we discuss below.

[12.2] Is the world classical? (EPR, Bell)

We would like to examine whether the world we live in is "classical" or not. The notion of classical world includes mainly two ingredients: (i) realism (ii) determinism. By realism we means that any quantity that can be measured is well defined even if we do not measure it in practice. By determinism we mean that the result of a measurement is determined in a definite way by the state of the system and by the measurement setup. We shall see later that quantum mechanics is not classical in both respects: In the case of spin 1/2 we cannot associate a definite value of $\hat{\sigma}_y$ for a spin which has been polarized in the $\hat{\sigma}_x$ direction. Moreover, if we measure the $\hat{\sigma}_y$ of a $\hat{\sigma}_x$ polarized spin, we get with equal probability ± 1 as the result.

In this section we would like to assume that our world is "classical". Also we would like to assume that interactions cannot travel faster than light. In some textbooks the latter is called "locality of the interactions" or "causality". It has been found by Bell that the two assumptions lead to an inequality that can be tested experimentally. It turns out from actual experiments that Bell's inequality are violated. This means that our world is either non-classical or else we have to assume that interactions can travel faster than light.

If the world is classical it follows that for any set of initial conditions a given measurement would yield a definite result. Whether or not we know how to predict or calculate the outcome of a possible measurement is not assumed. To be specific let us consider a particle of zero spin, which disintegrates into two particles going in opposite directions, each with spin 1/2. Let us assume that each spin is described by a set of state variables.

$$\begin{aligned} \text{state of particle A} &= x_1^A, x_2^A, ... \\ \text{state of particle B} &= x_1^B, x_2^B, ... \end{aligned} \tag{12.1}$$

The number of state variables might be very big, but it is assumed to be a finite set. Possibly we are not aware or not able to measure some of these "hidden" variables.

Since we possibly do not have total control over the disintegration, the emerging state of the two particles is described by a joint probability function $\rho\left(x_1^A, ..., x_1^B, ...\right)$. We assume that the particles do not affect each other after the disintegration ("causality" assumption). We measure the spin of each of the particles using a Stern-Gerlach apparatus. The measurement can yield either 1 or -1. For the first particle the measurement outcome will be denoted as a, and for the second particle it will be denoted as b. It is assumed that the outcomes a and b are determined in a deterministic fashion. Namely, given the state variables of the particle and the orientation θ of the apparatus we have

$$\begin{aligned} a &= f(\theta_A, x_1^A, x_2^A, ...) = \pm 1 \\ b &= f(\theta_B, x_1^B, x_2^B, ...) = \pm 1 \end{aligned} \tag{12.2}$$

where the function $f()$ is possibly very complicated. If we put the Stern-Gerlach machine in a different orientation then we will get different results:

$$\begin{aligned} a' &= f\left(\theta'_A, x_1^A, x_2^A, ...\right) = \pm 1 \\ b' &= f\left(\theta'_B, x_1^B, x_2^B, ...\right) = \pm 1 \end{aligned} \tag{12.3}$$

We have the following innocent identity:

$$ab + ab' + a'b - a'b' = \pm 2 \tag{12.4}$$

The proof is as follows: if $b = b'$ the sum is $\pm 2a$, while if $b = -b'$ the sum is $\pm 2a'$. Though this identity looks innocent, it is completely non trivial. It assumes both "reality" and "causality". The realism is reflected by the assumption that both a and a' have definite values, as implied by the function $f()$, even if we do not measure them. In the classical context it is not an issue whether there is a practical possibility to measure both a and a' at a single run of the experiment. As for the causality: it is reflected by assuming that a depends on θ_A but not on the distant setup parameter θ_B.

Let us assume that we have conducted this experiment many times. Since we have a joint probability distribution ρ, we can calculate average values, for instance:

$$\langle ab \rangle = \int \rho\left(x_1^A, ..., x_1^B, ...\right) f\left(\theta_A, x_1^A, ...\right) f\left(\theta_B, x_1^B, ...\right) \tag{12.5}$$

Thus we get that the following inequality should hold:

$$\left| \langle ab \rangle + \langle ab' \rangle + \langle a'b \rangle - \langle a'b' \rangle \right| \leq 2 \tag{12.6}$$

This is called Bell's inequality (in fact it is a variation of the original version). Let us see whether it is consistent with quantum mechanics. We assume that all the pairs are generated in a singlet (zero angular momentum) state. It is not difficult to calculate the expectation values. The result is

$$\langle ab \rangle = -\cos(\theta_A - \theta_B) \equiv C(\theta_A - \theta_B) \tag{12.7}$$

we have for example

$$C(0^o) = -1, \qquad C(45^o) = -\frac{1}{\sqrt{2}}, \qquad C(90^o) = 0, \qquad C(180^o) = +1. \tag{12.8}$$

If the world were classical the Bell's inequality would imply

$$|C(\theta_A - \theta_B) + C(\theta_A - \theta'_B) + C(\theta'_A - \theta_B) - C(\theta'_A - \theta'_B)| \leq 2 \tag{12.9}$$

Let us take $\theta_A = 0^o$ and $\theta_B = 45^o$ and $\theta'_A = 90^o$ and $\theta'_B = -45^o$. Assuming that quantum mechanics holds we get

$$\left|\left(-\frac{1}{\sqrt{2}}\right) + \left(-\frac{1}{\sqrt{2}}\right) + \left(-\frac{1}{\sqrt{2}}\right) - \left(+\frac{1}{\sqrt{2}}\right)\right| \;=\; 2\sqrt{2} \;>\; 2 \tag{12.10}$$

It turns out, on the basis of celebrated experiments that Nature has chosen to violate Bell's inequality. Furthermore it seems that the results of the experiments are consistent with the predictions of quantum mechanics. Assuming that we do not want to admit that interactions can travel faster than light it follows that our world is not classical.

In order to generalize the Bell inequality for other systems, it is worth noting that it is a variation of $\langle F \rangle < \sqrt{\langle F^2 \rangle}$. Above $F = (ab + ab' + a'b - a'b')$. The issue is that $\left\langle F^2 \right\rangle_{cl}$ provides a lower bound, while $\left\langle F^2 \right\rangle_{qm}$ is larger.

[12.3] Optional tests of realism

Mermin and Greenberger-Horne-Zeilinger have proposed optional tests for realism. The idea is to show that the *feasibility* of preparing some quantum states cannot be explained within the framework of a *classical* theory. We provide below two simple examples. The spin 1/2 mathematics that is required to understand these examples will be discussed in later lecture. What we need below is merely the following identities that express polarizations in the X and Y directions as a superposition of polarizations in the Z direction:

$$|x\rangle = \frac{1}{\sqrt{2}}\left(|z\rangle + |\bar{z}\rangle\right) \tag{12.11}$$

$$|\bar{x}\rangle = \frac{1}{\sqrt{2}}\left(|z\rangle - |\bar{z}\rangle\right) \tag{12.12}$$

$$|y\rangle = \frac{1}{\sqrt{2}}\left(|z\rangle + i|\bar{z}\rangle\right) \tag{12.13}$$

$$|\bar{y}\rangle = \frac{1}{\sqrt{2}}\left(|z\rangle - i|\bar{z}\rangle\right) \tag{12.14}$$

We use the notations $|z\rangle$ and $|\bar{z}\rangle$ for denoting "spin up" and "spin down" in Z polarization measurement, and similar convection for polarization measurement in the other optional directions X and Y.

Three spin example.– Consider 3 spins that are prepared in the following superposition state:

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\uparrow\uparrow\rangle - |\downarrow\downarrow\downarrow\rangle\right) \equiv \frac{1}{\sqrt{2}}\left(|zzz\rangle - |\bar{z}\bar{z}\bar{z}\rangle\right) \tag{12.15}$$

If we measure the polarization of 3 spins we get $a = \pm 1$ and $b = \pm 1$ and $c = \pm 1$, and the product would be $C = abc = \pm 1$. If the the measurement is in the ZZZ basis the result might be either $C_{ZZZ} = +1$ or $C_{ZZZ} = -1$

with equal probabilities. But optionally we can perform an XXX measurement or XYY, or YXY, or YYX measurement. If for example we perform XYY measurement it is useful to write the state in the XYY basis:

$$|\psi\rangle \;=\; \frac{1}{2}\left(|\bar{x}y\bar{y}\rangle + |\bar{x}\bar{y}y\rangle + |x\bar{y}\bar{y}\rangle + |xyy\rangle\right) \tag{12.16}$$

We see that the product of polarization is always $C_{XYY} = +1$. Similarly one can show that $C_{YXY} = +1$ and $C_{YYX} = +1$. If the world were classical we could predict the result of an XXX measurement:

$$C_{XXX} \;=\; a_x b_x c_x \;=\; a_x\, b_x\, c_x\, a_y^2\, b_y^2\, c_y^2 \;=\; C_{XYY}C_{YXY}C_{YYX} \;=\; 1 \tag{12.17}$$

But quantum theory predicts a contradicting result. To see what is the expected result we write the state in the XXX basis:

$$|\psi\rangle \;=\; \frac{1}{2}\left(|\bar{x}xx\rangle + |x\bar{x}x\rangle + |xx\bar{x}\rangle + |\bar{x}\bar{x}\bar{x}\rangle\right) \tag{12.18}$$

We see that the product of polarization is always $C_{XXX} = -1$. Thus, the experimental feasibility of preparing such quantum state contradicts classical realism.

Two spin example.– Consider 2 spins that are prepared in the following superposition state:

$$|\psi\rangle \;=\; \frac{1}{\sqrt{3}}\left(|z\bar{z}\rangle + |\bar{z}z\rangle - |\bar{z}\bar{z}\rangle\right) \tag{12.19}$$

$$\;=\; \frac{1}{\sqrt{6}}\left(|zx\rangle - |z\bar{x}\rangle + 2|\bar{z}\bar{x}\rangle\right) \tag{12.20}$$

$$\;=\; \frac{1}{\sqrt{6}}\left(|xz\rangle - |\bar{x}z\rangle + 2|\bar{x}\bar{z}\rangle\right) \tag{12.21}$$

$$\;=\; \frac{1}{\sqrt{12}}\left(|xx\rangle + |x\bar{x}\rangle + |\bar{x}x\rangle - 3|\bar{x}\bar{x}\rangle\right) \tag{12.22}$$

Above we wrote the state in the optional bases ZZ and ZX and XZ and XX. By inspection we see the following:

(1) The ZZ measurement result $|zz\rangle$ is impossible.
(2) The ZX measurement result $|\bar{z}x\rangle$ is impossible.
(3) The XZ measurement result $|x\bar{z}\rangle$ is impossible.
(4) All XX measurement results are possible with finite probability.

We now realize that in a classical reality observation (4) is in contradiction with observations (1-3). The argument is as follow: in each run of the experiment the state $\vec{a} = (a_x, a_z)$ of the first particle is determined by some set of hidden variables. The same applies with regard to the $\vec{b} = (b_x, b_z)$ of the second particle. We can define a *joint* probability function $f\left(\vec{a}, \vec{b}\right)$ that gives the probabilities to have any of the 4×4 possibilities (irrespective of what we measure in practice). It is useful to draw a 4×4 truth table and to indicate all the possibilities that are not compatible with (1-3). Then it turns out that the remaining possibilities are all characterized by having $a_x = -1$ or $b_x = -1$. This means that in a classical reality the probability to measure $|xx\rangle$ is zero. This contradicts the quantum prediction (4). Thus, the experimental feasibility of preparing such quantum state contradicts classical realism.

[12.4] The singlet state and the Pauli principle

שני אלקטרונים יכולים לאכלס את אותו אורביטל כיוון שאחד יכול להתישב עם ספין up והשני עם ספין $down$. זה ניסוח מקובל בספרי לימוד אלמנטריים. מבחינה מתמטית הרעיון דורש הבהרה יותר רצינית. לכאורה משתמע שהמצב הקוונטי של שני האלקטרונים הוא

$$|\psi\rangle = |\uparrow\downarrow\rangle \tag{12.23}$$

במצב זה האלקטרונים לכאורה מקיימים את "האיסור של פאולי, וכמו כן לכאורה הספין הכולל הוא אפס. למעשה שתי הטענות אינן נכונות. בהרצאה הבאה נראה שאם היה מצב כזה, והיו מודדים את שני הספינים בכיוון ציר x, אז היתה הסתברות סופית למצוא אותם באותו כיוון. כך שלמעשה המצב שהוגדר לעיל אינו מקיים את האיסור של פאולי, והספין הכולל שהוא מייצג אינו אפס. לכאורה אפשר גם לחשוב על צרופים היפותתיים אחרים כגון $|\downarrow\uparrow\rangle,\ |\uparrow\rightarrow\rangle,\ |\rightarrow\leftarrow\rangle$ ולכולם יש את אותה "בעיה". יש כאן גם עניין "אסתטי": האלקטרונים הם חלקיקים זהים, בעוד שלמצבים האלה אין סימטריה מוגדרת. למעשה רק "מצב איכלוס" אחד שהוא בא בחשבון:

$$|\text{singlet}\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \tag{12.24}$$

מצב זה נקרא "סינגלט". במצב כזה הספין הכולל של שני האלקטרונים הוא אפס בכל כיוון מדידה. מכאן גם שהמצב הזה מקיים את האיסור של פאולי: לא ניתן למצוא את האלקטרונים עם ספינים באותו כיוון. יש להדגיש שמצב הסינגלט הוא אחד ויחיד: בכל שאר המצבים אפשר למצוא את האלקטרונים עם ספין כולל שונה מאפס. בנוסף, למצב הזה יש סימטריה מוגדרת: הוא אנטי-סימטרי ביחס להחלפת החלקיקים. בפורמאליזם המסורתי של המכניקה הקוונטית דורשים שפונקצית הגל של הפרמיונים תהיה "אנטי סימטרית": זה למעשה הניסוח המתמטי של "עקרון האיסור של פאולי".

נחדד את הרעיון הזה מבחינה מתמטית. נתונה מערכת של שני חלקיקים שלכל אחד יש ספין חצי. נרצה למצוא מצב שבו הספין הכולל של המערכת הוא $S = 0$ באופן ודאי)בהסתברות 100%(. להלן נוח להשתמש בסימון $|z\rangle,\ |\bar{z}\rangle$ במקום הסימון $|\uparrow\rangle,\ |\downarrow\rangle$ נתבונן תחילה במצב $|z\bar{z}\rangle$ מצב זה אינו מאופין בתנע זויתי כולל אפס. כדי להיווכח בכך נניח שאנו מודדים את הקיטוב בכיוון אופקי. נרשום את המצב בבסיס שמתאים למדידה כזו ונקבל:

$$|z\bar{z}\rangle = \frac{1}{2}\left(|xx\rangle - |x\bar{x}\rangle + |\bar{x}x\rangle - |\bar{x}\bar{x}\rangle\right) \tag{12.25}$$

אנו רואים ש הסתברות של 25% למדוד ספין כולל 1 והסתברות 25% למדוד ספין 1-. במילים אחרות, במדידה אופקית יש הסתברות של 50% למצוא את שני הספינים באותו כיוון. למעשה יש רק מצב אחד של המערכת שבו מתקבל ספין כולל אפס בכל מדידה אפשרית - זה מצב הסינגלט שהגדרנו קודם. נשים לב שמתקיים

$$|\text{singlet}\rangle = \frac{1}{\sqrt{2}}\left(|z\bar{z}\rangle - |\bar{z}z\rangle\right) = \frac{1}{\sqrt{2}}\left(|x\bar{x}\rangle - |\bar{x}x\rangle\right) = \frac{1}{\sqrt{2}}\left(|\theta\bar{\theta}\rangle - |\bar{\theta}\theta\rangle\right) \tag{12.26}$$

את השיוויון האחרון אפשר להוכיח בקלות באמצעות ההצבות

$$|\theta\rangle = \cos\frac{\theta}{2}|z\rangle + \sin\frac{\theta}{2}|\bar{z}\rangle \tag{12.27}$$

$$|\bar{\theta}\rangle = \sin\frac{\theta}{2}|z\rangle - \cos\frac{\theta}{2}|\bar{z}\rangle \tag{12.28}$$

השורה התחתונה היא שבמצב סינגלט, גם במדידה בכיוון אופקי, וגם בכל כיוון מדידה אחר, מתקבל ספין כולל אפס בהסתברות של 100%.

[12.5] Correlations of entangled spins

נתייחס לשני חלקיקים שהספינים שלהם במצב סינגלט. מצב כזה יכול למשל להתקבל בהתפרקות של חלקיק בעל ספין אפס לשני חלקיקים בעלי ספין חצי. בגלל חוק שימור התנע הזויתי הספין הכולל לאחר ההתפרקות יהיה עדיין אפס - כך שהמצב חייב להיות מצב סינגלט. במצב כזה יש קורלציה בין החלקיקים. אם מודדים את הקיטוב האופקי אז אם אחד מהספינים הוא up אז השני חייב להיות $down$. נתיחס עתה למקרה כללי יותר. נניח שאנו מודדים את הקיטוב של הספין הראשון בכיוון אנכי, ואת הקיטוב של הספין השני בכיוון θ נסמן את תוצאות המדידה בצורה הבאה:

$$a = \pm 1 \tag{12.29}$$

$$b = \pm 1 \tag{12.30}$$

נאפיין את הקורלציה בין שתי התוצאות באמצעות המכפלה $c = ab$. קל להשתכנע שמתקיים

$$\text{Probability}(c = -1) = \left|\cos\left(\frac{\theta}{2}\right)\right|^2 \tag{12.31}$$

$$\text{Probability}(c = +1) = \left|\sin\left(\frac{\theta}{2}\right)\right|^2 \tag{12.32}$$

מכאן אנו מקבלים את הקורלציה

$$C(\theta) = \langle c \rangle = \langle ab \rangle = -\cos(\theta) \tag{12.33}$$

נבדוק את התוצאה שקיבלנו עבור זוויות מעניינות:
עבור $\theta = 0$ נקבל כי $\langle ab \rangle = -1$. משמעות הדבר היא שאם $a = 1$ אז בוודאות $b = -1$.
עבור $\theta = \pi$ נקבל כי $\langle ab \rangle = +1$. משמעות הדבר היא שאם $a = 1$ אז בוודאות $b = 1$.
עבור $\theta = \pi/2$ נקבל כי $\langle ab \rangle = 0$. משמעות הדבר היא שאם $a = 1$ אז יש סיכוי שווה לקבל $b = 1$ או $b = -1$.