III. Classical and Quantum Frameworks

In this section we describe the basic ingredients of the mathematical framework of classical and quantum computing in a somewhat peculiar way. That is, we first introduce the frameworks for classical mechanics of point particles, then for classical mechanics formulated as transport theory, and finally for quantum mechanics. Having introduced these we then develop the notion of classical computing in parallel to the classical mechanics of point particles. Similarly, we return to probabilistic methods of computation, such as simulated annealing, and finally come to the basic rules of quantum computing.

This approach is somewhat unorthodox, as it does not start with Qubits as the generalizations in quantum computing of bits. We hope, however, that basic concepts are clearer this way.

III.1. Classical and Quantum Mechanics of Particles in Space

Classical Mechanics of Point Particles. Suppose we wish to describe the motion of $N \in \mathbb{N}$ point particles moving in space $\mathbb{R}^3$. The motion of each particle is described by its position $q \in \mathbb{R}^3$ and its momentum $p \in \mathbb{R}^3$ which is put together in a phase space coordinate $x = (q, p) \in \Omega^{(1)} := \mathbb{R}^3 \times \mathbb{R}^3$. Then $x_k(t) = \big(q_k(t), p_k(t)\big) \in \Omega^{(1)}$ denotes the phase space coordinates of the k^{th} particle at a given time $t \in \mathbb{R}$. Putting these coordinates together in one N-tuple, the vector

$$\underline{x}(t) \;:=\; \big(x_1(t),\, x_2(t),\, \ldots,\, x_N(t)\big) \;\in\; \Omega^{(N)} \;:=\; [\Omega^{(1)}]^N\,, \tag{III.1}$$

encodes the complete desciption of the particles' phase space coordinates which we call phase space configuration. For this reason $\Omega^{(N)}$ is called the **phase space** of the system of N particles.

Almost two hundred years ago, Hamilton obtained a formulation of Newtonian mechanics that determines the phase space configuration $\underline{x}(t) \in \Omega^{(N)}$ of the system at time $t \in \mathbb{R}$ under

the assumption that the configuration $\underline{x}(s) \in \Omega^{(N)}$ is known for some $s < t$. For the sake of simplicity, we assume that $t > 0 = s$. The **dynamics** of the particles is determined by the *Hamilton equations of motion* which is a system of first-order ordinary differential equations, namely,

$$\forall\, t > 0: \quad \dot{\underline{x}}(t) \;=\; H'[\underline{x}(t)]\,, \qquad \underline{x}(0) = \underline{x}_0\,, \tag{III.2}$$

where

$$H'[\underline{x}(t)] \;:=\; \Big(\nabla_{\underline{p}} H[\underline{x}(t)],\; -\nabla_{\underline{q}} H[\underline{x}(t)]\Big) \tag{III.3}$$

is a (symplectic) gradient of the *Hamiltonian function* $H \in C^1(\Omega^{(N)}; \Omega^{(N)})$ of the system whose precise form is immaterial for our purpose. We do not go into the fairly complicated theory of existence and uniqueness of the solutions of Hamilton's equation (III.2) of motion. Instead we assume these two properties by demanding that there exists a **flow** (map) $\Phi \in C^1(\mathbb{R} \times \Omega^{(N)}; \Omega^{(N)})$ such that

$$\forall\, (t, \underline{x}_0) \in \mathbb{R} \times \Omega^{(N)}: \quad \underline{x}(t) \;=\; \Phi_t(\underline{x}_0)\,. \tag{III.4}$$

That is, for any initial configuration $\underline{x}_0 \in \Omega^{(N)}$ of spatial positions and momenta, the system of N particles follows the **trajectory** $\big(t \mapsto \Phi_t(\underline{x}_0)\big) \in C^1(\mathbb{R}_0^+; \Omega^{(N)}$.

Suppose we are now given an **observable**, i.e., a measurable physical quantity represented by a bounded, measurable real function $A \in L^\infty(\Omega^{(N)}; \mathbb{R})$ on phase space. A good example to have in mind, although neither bounded nor real-valued, is the center of mass $A : \Omega^{(N)} \to \mathbb{R}^3$ of the particles defined for $\underline{x} = (q_1, \ldots, q_N, p_1, \ldots, p_N)$ as

$$A[\underline{x}] \;:=\; \frac{1}{N} \sum_{k=1}^{N} q_k\,. \tag{III.5}$$

The center of mass $A_t[\underline{x}_0] \in \mathbb{R}^3$ of the particles at time $t > 0$ with initial configuration $\underline{x}_0 \in \Omega^{(N)}$ is then given by

$$A_t[\underline{x}_0] \;:=\; A[\underline{x}(t)] \;=\; A[\Phi_t(\underline{x}_0)]\,. \tag{III.6}$$

This equation holds true for any observable $A \in L^\infty(\Omega^{(N)}; \mathbb{R})$ - not only for the center of mass.

Probabilistic Formulation of Classical Mechanics. We now broaden our perspective and describe the system's configurations at time t not by points $\underline{x}(t)$ in phase space but by *functions* (or, more precisely, probability distributions) *on phase space*. For this, we replace the initial configuration $\underline{x}_0 \in \Omega^{(N)}$ by an initial phase space density, i.e., a probability distribution $\rho_0 \in L^1(\Omega^{(N)}; \mathbb{R}_0^+)$, with $\int_{\Omega^{(N)}} \rho_0(\underline{x})\, d\underline{x} = 1$, where $d\underline{x} = d^N q\, d^N p$ is Lebesgue measure on phase space $\Omega^{(N)}$. The physical state of the system is now represented by the function $\rho_0 : \Omega^{(N)} \to \mathbb{R}_0^+$, not the point $\underline{x}_0 \in \Omega^{(N)}$. As each of the points $\underline{x}_0 \in \Omega^{(N)}$, considered a

potential initial configuration, follows its trajectory, the state of the system at time $t > 0$ is then given by

$$\rho_t \;=\; \rho_0 \circ \Phi_{-t}\,. \tag{III.7}$$

(Note the minus sign in the time variable.) Since the flow Φ_t leaves the measure on phase space invariant, $\rho_t \in L^1(\Omega^{(N)};\mathbb{R}_0^+)$ is a probability distribution, too. We now evaluate an observable $A \in L^\infty(\Omega^{(N)};\mathbb{R})$ at time $t \geq 0$. Since ρ_t is a probability distribution, this is actually an *expectation value* and is given by

$$\langle A\rangle_{\rho_t} \;:=\; \int_{\Omega^{(N)}} A[\underline{x}]\;\rho_t(\underline{x})\;d\underline{x} \;=\; \int_{\Omega^{(N)}} A[\underline{x}]\;\rho_0[\Phi_{-t}(\underline{x})]\;d\underline{x}\,. \tag{III.8}$$

Note that our first description in terms of phase space points is contained in this larger framework if, slightly more generally, we allow ρ_0 to be a probabilty measure -not necessarily an integrable function- and assume $A \in C(\Omega^{(N)};\mathbb{R})$ to be continuous. Namely, choosing $\rho_0(\underline{x}) := \delta(\underline{x} - \underline{x}_0)$, Eq. (III.8) yields

$$\langle A\rangle_{\rho_t} \;=\; \int_{\Omega^{(N)}} A[\underline{x}]\;\delta[\Phi_{-t}(\underline{x}) - \underline{x}_0]\;d\underline{x} \;=\; \int_{\Omega^{(N)}} A[\underline{x}]\;\delta[\underline{x} - \Phi_t(\underline{x}_0)]\;d\underline{x} \;=\; A[\Phi_t(\underline{x}_0)]\,, \tag{III.9}$$

indeed. This computation also illustrates the necessity of the minus sign in (III.7).

The formulation in terms of functions (here: probability distributions) on phase space has two decisive advantages.

- The first is that the (expectation) value $\langle A\rangle_{\rho_t}$ of an observable is a *linear functional* of both the observable $A \in L^\infty(\Omega^{(N)};\mathbb{R})$ and the function $\rho_t \in L^1(\Omega^{(N)};\mathbb{R}_0^+)$. More precisely, for all functions $\rho \in L^1(\Omega^{(N)})$, the expectation value

$$\big(A \mapsto \langle A\rangle_\rho\big) \;\in\; \big[L^\infty(\Omega^{(N)})\big]^* \tag{III.10}$$

  defines a continuous linear functional on $L^\infty(\Omega^{(N)})$. Conversely, if $\alpha \in (0,1)$ and $\rho, \widetilde{\rho} \in L^1(\Omega^{(N)};\mathbb{R}_0^+)$ are two probability densities on $\Omega^{(N)}$ then so is $[\alpha\rho + (1-\alpha)\widetilde{\rho}] \in L^1(\Omega^{(N)};\mathbb{R}_0^+)$, and for all observables $A \in L^\infty(\Omega^{(N)})$, the expectation value of $\alpha\rho + (1-\alpha)\widetilde{\rho}$ is given by

$$\langle A\rangle_{\alpha\rho+(1-\alpha)\widetilde{\rho}} \;=\; \alpha\,\langle A\rangle_\rho \;+\; (1-\alpha)\,\langle A\rangle_{\widetilde{\rho}}\,. \tag{III.11}$$

- The second, more important, advantage is that the form $\rho_t = \rho_0 \circ \Phi_{-t}$ of the solution of the time evolution in terms of a flow map is special and rather rigid. Most of the important evolution equations in physics and technology other than classical mechanics do not posess solutions of this form. We illustrate this argument on the example of the heat equation in $\mathbb{R}^3$. Given an initial temperature profile $\rho_0 \in L^1(\mathbb{R}^3;\mathbb{R}_0^+)$, the temperature profile at time $t > 0$ is the unique solution u_t of the heat equation

$$\forall\, t > 0,\; x \in \mathbb{R}^3 : \quad \dot{u}_t(x) \;=\; \Delta_x u_t(x)\,, \qquad u_0(x) = \rho_0(x)\,, \tag{III.12}$$

which can be explicitly computed by convolving the initial profile with the heat kernel,

$$\forall\, t > 0,\ x \in \mathbb{R}^3 : \quad u_t(x) \;=\; \int e^{-(x-y)^2/(4t)}\, \rho_0(y)\, \frac{d^3y}{(4\pi t)^{3/2}}\,. \tag{III.13}$$

There exists no flow map such that u_t could be written in the form $u_t(x) = \rho_0[\Phi_t(x)]$, for all $x \in \mathbb{R}^3$. (Nevertheless, the Ansatz $u_t(x) = \rho_0[\Phi_t(x)]$ is a useful method known as the *method of characteristics* in PDE theory to construct approximate solutions for small times.)

The corresponding generalization of Eq. (III.7) results from assuming that, given $t > 0$, there exists a conditional probability distribution $p_t : \Omega^{(N)} \times \Omega^{(N)} \to \mathbb{R}_0^+$ such that

$$\forall\, \underline{x} \in \Omega^{(N)} : \qquad \rho_t(\underline{x}) \;=\; \int_{\Omega^{(N)}} p_t(\underline{x}|\underline{y})\, \rho_0(\underline{y})\, d\underline{y}\,. \tag{III.14}$$

The requirement that p_t be a conditional probability distribution reads

$$\forall\, \underline{y} \in \Omega^{(N)} : \qquad \int_{\Omega^{(N)}} p_t(\underline{x}|\underline{y})\, d\underline{x} \;=\; 1\,, \tag{III.15}$$

which ensures by Fubini that

$$\int \rho_t(\underline{x})\, d\underline{x} \;=\; \iint p_t(\underline{x}|\underline{y})\, \rho_0(\underline{y})\, d\underline{x}\, d\underline{y} \;=\; \int \rho_0(\underline{y})\, d\underline{y}\,, \tag{III.16}$$

i.e., if $\rho_0 \in L^1(\Omega^{(N)}; \mathbb{R}_0^+)$ is a probability distribution then so is $\rho_t \in L^1(\Omega^{(N)}; \mathbb{R}_0^+)$. (Note that we are generous about measure-theoretic details such as (III.15) that actually is only required almost everywhere in $\Omega^{(N)}$.)

We illustrate the formulation (III.14) by two examples.

- The original motion of N point particles can be formulated as in (III.14) if we set

  $$p_t(\underline{x}|\underline{y}) \;:=\; \delta[\underline{x} - \Phi_t(\underline{y})]\,, \tag{III.17}$$

  since, with this choice of p_t, we obtain

  $$\rho_t(\underline{x}) \;=\; \int_{\Omega^{(N)}} \delta[\underline{x} - \Phi_t(\underline{y})]\, \rho_0(\underline{y})\, d\underline{y} \;=\; \rho_0[\Phi_{-t}(\underline{x})]\,. \tag{III.18}$$

- The heat kernel $p_t(x|y) = (4\pi t)^{-3/2} \exp[-(x-y)^2/(4t)]$ used in (III.13) to solve the heat equation is a second, typical example for such conditional probability distribution. Indeed, for all $y \in \mathbb{R}^3$ and $t > 0$,

  $$\int_{\mathbb{R}^3} p_t(x|y)\, d^3x \;=\; \int_{\mathbb{R}^3} \frac{e^{-(x-y)^2/(4t)}\, d^3x}{(4\pi t)^{3/2}} \;=\; \int_{\mathbb{R}^3} \frac{e^{-x^2}\, d^3x}{\pi^{3/2}} \;=\; 1\,. \tag{III.19}$$

Compared to quantum mechanics, the probablilistic formulation of classical mechanics has the disadvantage that, unless we are in a special case like (III.17), the dynamics is irreversible: Given ρ_0, we can compute ρ_t for $t > 0$, but given $t > 0$ and ρ_t, the reconstruction of the initial data ρ_0 is a rather complicated and, in fact, in many cases impossible. For instance, if u_t is the solution of the heat equation according to (III.13) then it is not hard to see that $u_t \in C^\infty(\mathbb{R}^3)$ is smooth. Consequently, if $t > 0$ and u_t is lacking this high regularity, there is no initial datum u_0 for which u_t is the solution of the heat equation at time t.

Quantum Mechanics. Similar to the probabilistic description of mechanics, we do not represent the state of the system of N point particles in quantum mechanics by a configuration $\underline{x}(t)$ in phase space, but by a complex-valued function ψ_t on the space of configurations $\Omega^{(N)}$. We now go through the construction step by step.

- As opposed to classical mechanics, the configuration space of a quantum mechanical particle contains only positions, not momenta. In a first step the state of a particle at time $t \in \mathbb{R}$ is represented by a complex-valued, square-integrable function $\psi_t : \Omega^{(1)} \to \mathbb{C}$ of the particle's position $x \in \Omega^{(1)} := \mathbb{R}^3$. Thanks to their square-integrability, these functions are elements $\psi_t \in \mathfrak{h}$ of the Hilbert space $\mathfrak{h} := L^2(\mathbb{R}^3)$.
- We further assume ψ_t to be normalized, i.e.,

$$\|\psi_t\|_2^2 \;=\; \int_{\Omega^{(1)}} |\psi_t(\underline{x})|^2 \, d\underline{x} \;=\; 1\,, \tag{III.20}$$

  so that the square of the absolute value $|\psi_t|^2 : \mathbb{R}^3 \to \mathbb{R}_0^+$ allows for the intepretation to be the probability distribution of the particle at time t. I.e., $P_t(A) := \int_A |\psi_t(x)|^2 \, d^3x$ is the probability to find the particle in a (measurable) subset $A \subseteq \Omega^{(1)}$ of its configuration space $\Omega^{(1)}$.
- Similarly, the configuration space $\Omega^{(N)} := [\Omega^{(1)}]^N = (\mathbb{R}^3)^N$ of N quantum mechanical particles contains N positions in $\Omega^{(1)}$ (and no momenta). The state of the N-particle system at time $t \in \mathbb{R}$ is represented by a complex-valued, square-integrable function $\Psi_t : \Omega^{(N)} \to \mathbb{C}$ of the N-particle configurations $\underline{x} \in \Omega^{(N)}$. Again its square-integrability ensures that this function $\Psi_t \in \mathfrak{H}^{(N)}$ belongs to the Hilbert space $\mathfrak{H}^{(N)} := L^2(\Omega^{(N)})$.
- As a mathematical fact, if $\Omega = \Omega_1 \times \Omega_2 = \{(x_1, x_2) |\; x_1 \in \Omega_1, x_2 \in \Omega_2\}$ is the cartesian product of two sets Ω_1 and Ω_2 then $L^2(\Omega)$ is isomorphic to the tensor product $L^2(\Omega_1) \otimes L^2(\Omega_1)$.
- This can be generalized to N factors: If $(\Omega_n, \mathfrak{A}_n, \mu_n)$ are measure spaces, for all $n \in \mathbb{Z}_1^N$, and

$$\Omega \;=\; \Omega_1 \times \Omega_2 \times \cdots \times \Omega_N \;=\; \big\{(x_1, x_2, \ldots, x_N) \;\big|\; x_1 \in \Omega_1, x_N \in \Omega_N\big\} \tag{III.21}$$

  is their cartesian product, then

$$L^2(\Omega) \;=\; L^2(\Omega_1) \otimes L^2(\Omega_2) \otimes \cdots \otimes L^2(\Omega_N)\,. \tag{III.22}$$

The dynamics of the N-particle system is given by the Schrödinger equation,

$$\forall\, t \in \mathbb{R}: \quad \dot{\psi}_t \;=\; -iH\psi_t\,, \qquad \psi_0 \in \mathcal{H}^{(N)}\,. \tag{III.23}$$

Here, $H = H^*$ is the self-adjoint Hamiltonian operator acting on $\mathcal{H}^{(N)}$) (Very often H is actually an unbounded operator, but we ignore the mathematical complication that comes about with this unboundedness.) Thanks to its self-adjointness, H generates a one-parameter

group $(U_t)_{t\in\mathbb{R}} \subseteq \mathcal{U}(\mathcal{H}^{(N)})$ of unitary operators which is frequently called **propagator**, written as $\boldsymbol{U_t =: e^{-itH}}$. The unique solution of the Schrödinger equation is given by

$$\forall\, t \in \mathbb{R}: \qquad \psi_t \;=\; U_t\, \psi_0\,. \tag{III.24}$$

The Schrödinger equation can also be formulated as an evolution equation for the propagator, i.e.,

$$\forall\, t \in \mathbb{R}: \qquad \dot{U}_t \;=\; -iHU_t\,, \qquad U_0 = \mathbf{1}_{\mathcal{H}^{(N)}}\,. \tag{III.25}$$

Note that the quantum evolution (III.24) is perfectly reversible, namely, $\psi_0 = U_t^*\psi_t$, since $U_t^{-1} = U_t^*$, as U_t is unitary.

Observables in quantum mechanics are represented by self-adjoint operators $A = A^* \in \mathcal{B}(\mathcal{H}^{(N)})$. In fact, the Hamiltonian H is an observable, too, namely the system's energy. While H is an unbounded operator, we may always assume w.l.o.g. a given observable $A = A^* \in \mathcal{B}(\mathcal{H}^{(N)})$ to be bounded. Its expectation value at time $t \in \mathbb{R}$ is defined to be the diagonal matrix element

$$\forall\, t \in \mathbb{R}: \qquad \langle A_t\rangle_{\psi_0} \;=\; \langle\psi_0|A_t\psi_0\rangle \;:=\; \langle\psi_t|A\psi_t\rangle \;=\; \langle\psi_0|(U_t^*AU_t)\,\psi_0\rangle\,. \tag{III.26}$$

Note that this implies that $A_t = U_t^*AU_t$. Hence, using (III.25), we obtain the Heisenberg equation of motion

$$\forall\, t \in \mathbb{R}: \qquad \dot{A}_t \;=\; -i\big[H\,,\,A_t\big]\,, \tag{III.27}$$

which is actually equivalent to the Schrödinger equation.

Embedded Systems in Quantum Mechanics and Density Matrices Summarizing the framework of quantum mechanics presented so far, we note that states of a physical system S at time $t \in \mathbb{R}$ are represented by vectors $\psi_t \in \mathcal{H}_S$ in the system's Hilbert space $\mathcal{H}_S$, which typically is the space of complex-valued, square-integrable functions of the (classical spatial) coordinates $x \in \Omega_S$ of the system, i.e., $\mathcal{H}_S = L^2(\Omega_S)$.

Suppose now that S_1 is a physical system with coordinate space Ω_1 and quantum mechanical states in $\mathcal{H}_1 = L^2(\Omega_1)$. If S_1 is actually a subsystem of a larger total system S_{12} containing another subsystem S_2, besides S_1, with classical coordinates in Ω_2 then the coordinate space of the total system S_{12} is naturally $\Omega_{12} = \Omega_1 \times \Omega_2$, and the corresponding Hilbert space of states in S_{12} is $L^2(\Omega_{12}) = L^2(\Omega_1) \otimes L^2(\Omega_2)$.

More generally, tensor products appear in quantum mechanics whenever we have two physical subsystems S_1 and S_2 and the total system S_{12} consists of these two subsystems. If the states of S_1 and S_2 are vectors in a Hilbert space $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively, then the states of the total system are vectors in their tensor product $\mathcal{H}_{12} = \mathcal{H}_1 \otimes \mathcal{H}_2$.

Remarks and Examples.

- If S_1 represents an electron and S_2 represents a proton then $\Omega_1 = \Omega_2 = \mathbb{R}^3 \times \{\uparrow, \downarrow\}$ and $\mathcal{H}_1 = \mathcal{H}_2 = L^2(\mathbb{R}^3 \times \{\uparrow, \downarrow\})$. The total system S_{12} contains an electron and a proton and may be considered a hydrogen atom; its Hilbert space is $\mathcal{H}_{12} = \mathcal{H}_1 \otimes \mathcal{H}_2 = L^2(\mathbb{R}^3 \times \{\uparrow, \downarrow\}) \otimes L^2(\mathbb{R}^3 \times \{\uparrow, \downarrow\})$.
- If S_3 is the quantized photon field then its Hilbert space is the photon Fock space $\mathcal{F}(\mathfrak{h})$ over the one-photon Hilbert space $\mathfrak{h} = \left\{ f \in L^2(\mathbb{R}^3; \mathbb{R}^3) \,\middle|\, \forall\, \vec{k} \in \mathbb{R}^3 \setminus \{\vec{0}\} :\ \hat{f}(\vec{k}) \perp \vec{k} \right\}$ of square-integrable, divergent-free vector fields (Coulomb gauge).
- If the total system S_{123} consists of a hydrogen atom and the quantized radiation field then its Hilbert space is $\mathcal{H}_{123} = \mathcal{H}_{12} \otimes \mathcal{F}(\mathfrak{h})$.

We now consider a system S_1 whose states are represented by normalized vectors ψ_1 in a Hilbert space $\mathcal{H}_1$. We wish to account for the possibility that S_1 is a subsystem of a larger system S_{12} whose states are represented by normalized vectors Ψ_{12} in a Hilbert space $\mathcal{H}_{12} = \mathcal{H}_1 \otimes \mathcal{H}_2$, where $\mathcal{H}_2$ is the Hilbert space for the other constituent of S_{12}, namely, a subsystem S_2. For definiteness, we assume both $\mathcal{H}_1$ and $\mathcal{H}_2$ to be infinite dimensional.

If $A_1 = A_1^* \in \mathcal{B}(\mathcal{H}_1)$ is an observable of the system S_1 then its expectation value in the state $\Psi_{12} \in \mathcal{H}_{12}$ is given by

$$\langle A_1 \rangle_{\Psi_{12}} \;=\; \langle \Psi_{12} | (A_1 \otimes \mathbf{1}_2) \Psi_{12} \rangle_{12} \,. \tag{III.28}$$

This expectation value is of the form $\langle A_1 \rangle_{\Psi_{12}} = \langle \psi_1 | A_1 \psi_1 \rangle_{\mathcal{H}_1}$, for some $\psi_1 \in \mathcal{H}_1$ if, and only if, $\Psi_{12} = \psi_1 \otimes \psi_2$. This is, however, unphysical because it is equivalent to assuming the two subsystems S_1 and S_2 to be independent of each other and in absence of any interaction between them.

Now we suppose that $\Psi \in \mathcal{H}_{12}$ is an arbitray normalized vector. We define a linear operator $\rho_1 \in \mathcal{B}(\mathcal{H}_1)$ by

$$\big\langle f \big|\, \rho_1\, f' \big\rangle_1 \;:=\; \sum_{n=1}^{\infty} \Big\langle f \otimes g_n \Big| \big(|\Psi\rangle\langle\Psi| \big)\, f' \otimes g_n \Big\rangle_{12} \;=\; \sum_{n=1}^{\infty} \langle f \otimes g_n | \Psi \rangle_{12}\, \langle \Psi | f' \otimes g_n \rangle_{12} \,, \tag{III.29}$$

where $\{g_n\}_{n=1}^{\infty} \subseteq \mathcal{H}_2$ is an arbitrary ONB. A simple application of the Cauchy-Schwarz inequality shows that (III.29) is convergent and independent of the ONB $\{g_n\}_{n=1}^{\infty}$. Obviously, $\rho_1 \geq 0$ is positive. If $\{f_m\}_{m=1}^{\infty} \subseteq \mathcal{H}_1$ is an ONB then

$$\sum_{m=1}^{\infty} \big\langle f_m \big|\, \rho_1\, f_m \big\rangle_1 \;:=\; \sum_{m,n=1}^{\infty} \Big\langle f_m \otimes g_n \Big| \big(|\Psi\rangle\langle\Psi| \big)\, f_m \otimes g_n \Big\rangle_{12} \;=\; \|\Psi\|^2 \;=\; 1 \,. \tag{III.30}$$

Hence, $\rho_1 \in \boldsymbol{\mathcal{DM}(\mathcal{H}_1)}$ is a **density matrix**, i.e., a positive trace class operator on a Hilbert space $\mathcal{H}$ of trace one,

$$\mathcal{DM}(\mathcal{H}) \;:=\; \Big\{ \rho \in \mathcal{L}^1(\mathcal{H}) \;\Big|\; \rho \geq 0\,,\ \mathrm{Tr}(\rho) = 1 \Big\} \,. \tag{III.31}$$

We observe that $\mathcal{DM}(\mathcal{H}) \subseteq \mathcal{L}^1(\mathcal{H})$ is closed and convex. In fact, $\mathcal{DM}(\mathcal{H})$ is the convex hull of all pure states, i.e., density matrices of the form $|\psi\rangle\langle\psi|$, with ψ normalized.

Moreover, if $\rho_{12} \in \mathcal{DM}(\mathcal{H}_{12})$ is a density matrix of the total system S_{12} on the Hilbert space $\mathcal{H}_{12} = \mathcal{H}_1 \otimes \mathcal{H}_2$ and

$$\sum_{m=1}^{\infty} \big\langle f_m \big|\, \rho_1\, f_m \big\rangle_1 \;:=\; \sum_{m,n=1}^{\infty} \Big\langle f_m \otimes g_n \Big| \rho_{12}\, (f_m \otimes g_n) \Big\rangle_{12}\,, \tag{III.32}$$

then $\rho_1 \in \mathcal{DM}(\mathcal{H}_1)$ is a density matrix for the subsystem S_1 of S_{12}.

For this reason we replace wave functions by density matrices and represent states of physical systems by the latter, henceforth. If $A = A^* \in \mathcal{B}(\mathcal{H})$ is an observable of a system in a state represented by a density matrix $\rho \in \mathcal{DM}(\mathcal{H})$ on a Hilbert space $\mathcal{H}$, then its expectation value is given by

$$\langle A \rangle_\rho \;:=\; \mathrm{Tr}(\rho\, A)\,. \tag{III.33}$$

If the density matrix $\rho_t \in \mathcal{DM}(\mathcal{H})$ at time $t \in \mathbb{R}$ is the pure state $\rho_t = |\psi_t\rangle\langle\psi_t|$, then ρ_t results from the initial value ρ_0 by conjugation by the unitary propagator $U_t \in \mathcal{U}(\mathcal{H})$ of (III.24)-(III.25), i.e.,

$$\rho_t \;=\; |U_t\psi_0\rangle\langle U_t\psi_0| \;=\; U_t\, |\psi_0\rangle\langle\psi_0|\, U_t^* \;=\; U_t\, \rho_0\, U_t^* \tag{III.34}$$

Taking convex combinations of such pure states, we derive the dynamical law for a general density matrix $\rho_t \in \mathcal{DM}(\mathcal{H})$ representing the state of the system at time $t \in \mathbb{R}$, given its value $\rho_0 \in \mathcal{DM}(\mathcal{H})$ at $t = 0$,

$$\rho_t \;=\; U_t\, \rho_0\, U_t^*\,, \quad \dot\rho_t \;=\; -i[H, \rho_t]\,. \tag{III.35}$$

It is interesting to note that the equation of motion $\dot\rho_t \;=\; -i[H, \rho_t]$ is (potentially) easier to solve than the Schrödinger equation $\dot\psi_t \;=\; -iH\psi_t$, because the former does not follow oscillations of the phase in ψ_t anymore. Namely, for any choice of $\theta : \mathbb{R} \to \mathbb{R}$ and with $\psi_t^{(\theta)} := e^{i\theta(t)}\psi_t$, the density matrix $\rho_t = |\psi_t^{(\theta)}\rangle\langle\psi_t^{(\theta)}|$ is independent of θ.[1]

Another reason that lets density matrices appear superior to wave functions is that, while a linear combination of wave functions is again a wave function, its normalization is not preserved, in general. In contrast, the set $\mathcal{DM}(\mathcal{H}) \subseteq \mathcal{L}^1(\mathcal{H})$ of density matrices over a Hilbert space $\mathcal{H}$ is convex (and closed). So, given two density matrices $\rho_0, \rho_1 \in \mathcal{DM}(\mathcal{H})$ and $\alpha \in [0, 1]$, the operator $\rho_\alpha := (1-\alpha)\rho_0 + \alpha\rho_1 \in \mathcal{DM}(\mathcal{H})$ is a density matrix, as well.

Remarks and Examples. We exemplify this on a single qubit, i.e., $\mathcal{H} = \mathbb{C}^2$. We analyze the space $\mathcal{SA}(\mathcal{H}) \subseteq \mathcal{B}(\mathcal{H}) \cong \mathbb{C}^{2\times 2}$ by first observing that if $A \in \mathcal{B}(\mathcal{H})$ is given by

$$A \;=\; \begin{pmatrix} a & b \\ c & d \end{pmatrix}\,, \tag{III.36}$$

[1] I thank Thierry Paul for sharing this observation with me.

for some $a, b, c, d \in \mathbb{C}$ then $A = A^*$ iff $a, d \in \mathbb{R}$ and $c = \bar{b}$. So, any self-adjoint complex 2×2 matrix can be written as

$$\begin{aligned} A \;&=\; \begin{pmatrix} \alpha + \delta & \beta + i\gamma \\ \beta - i\gamma & \alpha - \delta \end{pmatrix} \\ &= \alpha\, \mathbf{1} \;+\; \beta\, \sigma^{(1)} \;+\; \gamma\, \sigma^{(2)} \;+\; \delta\, \sigma^{(3)}\,, \end{aligned} \tag{III.37}$$

for unique numbers $\alpha, \beta, \gamma, \delta \in \mathbb{R}$, where $\mathbf{1} \in \mathbb{C}^{2\times 2}$ is the unit and $\sigma^{(1)}, \sigma^{(2)}, \sigma^{(3)} \in \mathbb{C}^{2\times 2}$ are the self-adjoint Pauli matrices defined as

$$\mathbf{1} \;=\; \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\,,\;\; \sigma^{(1)} \;=\; \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\,,\;\; \sigma^{(2)} \;=\; \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}\,,\;\; \sigma^{(3)} \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,. \tag{III.38}$$

It is convenient to equip the real vector space $\mathcal{SA}(\mathbb{C}^2)$ with the scalar product $\langle A, B\rangle := \mathrm{Tr}(AB)$ which makes it a real Hilbert space $\big(\mathcal{SA}(\mathbb{C}^2), \langle\cdot,\cdot\rangle\big)$. Using the fact that

$$\sigma^{(j)}\, \sigma^{(k)} \;=\; \delta_{j,k} \cdot \mathbf{1} \;+\; i \sum_{\ell=1}^{3} \varepsilon_{jk\ell}\, \sigma^{(\ell)}\,, \tag{III.39}$$

where the *totally antisymmetric symbol* $\varepsilon_{jk\ell}$ is defined by

$$\forall\, \{j,k,\ell\} \subseteq \{1,2,3\}: \quad \varepsilon_{jk\ell} \;:=\; \begin{cases} \mathrm{sgn} \begin{pmatrix} 1 & 2 & 3 \\ j & k & \ell \end{pmatrix}\,, & \text{for } \{j,k,\ell\} \;=\; \{1,2,3\}\,, \\ 0, & \text{for } \{j,k,\ell\} \;\neq\; \{1,2,3\}\,, \end{cases} \tag{III.40}$$

it is easy to check that

$$\big\{ \tfrac{1}{\sqrt{2}}\, \mathbf{1},\; \tfrac{1}{\sqrt{2}}\, \sigma^{(1)},\; \tfrac{1}{\sqrt{2}}\, \sigma^{(2)},\; \tfrac{1}{\sqrt{2}}\, \sigma^{(3)} \big\} \;\subseteq\; \mathcal{SA}(\mathbb{C}^2) \tag{III.41}$$

is orthonormal and thus an ONB, since $\dim_{\mathbb{R}}[\mathcal{SA}(\mathbb{C}^2)] = 4$. This implies that

$$\forall\, A \in \mathcal{SA}(\mathbb{C}^2): \quad A \;= \tfrac{1}{2}\mathrm{Tr}(A)\, \mathbf{1} \;+\; \tfrac{1}{2} \vec{v}_A \cdot \vec{\sigma}\,, \tag{III.42}$$

where $\vec{\sigma} = \big(\sigma^{(1)}, \sigma^{(2)}, \sigma^{(3)}\big)^t$, $\vec{v}_A = \big(v_A^{(1)}, v_A^{(2)}, v_A^{(3)}\big)^t$, and

$$\mathrm{Tr}(A) \;=\; \langle \mathbf{1}, A\rangle \quad \text{and} \quad v_A^{(j)} \;=\; \langle \sigma^{(j)}, A\rangle \;=\; \mathrm{Tr}\big(\sigma^{(j)}\, A\big)\,. \tag{III.43}$$

Moreover, from (III.37) we see that

$$\det(A) \;=\; \alpha^2 - \beta^2 - \gamma^2 - \delta^2 \;=\; \frac{1}{4}\Big([\mathrm{Tr}(A)]^2 - |\vec{v}_A|^2 \Big)\,. \tag{III.44}$$

Specifically, if $\rho \in \mathcal{DM}(\mathbb{C}^2)$ is a density matrix then it is positive, and therefore its determinant is nonnegative. Thus $|\vec{v}_\rho| \leq \mathrm{Tr}(\rho) = 1$, i.e., $\vec{v}_\rho \in \overline{B(0,1)} \subseteq \mathbb{R}^3$ is a vector of length less or equal to one in three-dimensional Euclidean space. Moreover, ρ has the eigenvalues

$\lambda, 1-\lambda \in [0,1]$. It follows that $\lambda(1-\lambda) = \det(\rho) = (1-\vec{v}_\rho^{\,2})/4$ which, in turn, is equivalent to

$$\sigma(\rho) \;=\; \{\lambda,\, 1-\lambda\} \;=\; \Big\{\tfrac{1}{2}\big(1-|\vec{v}_\rho|\big)\,,\; \tfrac{1}{2}\big(1+|\vec{v}_\rho|\big)\Big\}\,. \tag{III.45}$$

In summary, it follows that

$$\mathcal{DM}(\mathbb{C}^2) \;=\; \Big\{\tfrac{1}{2}\big(\mathbf{1} \;+\; \vec{v}\cdot\vec{\sigma}\big)\;\Big|\vec{v}\in\mathbb{R}^3\,,\; |\vec{v}|_{\rm eucl}\leq 1\Big\}\,, \tag{III.46}$$

i.e., the convex set of density matrices on $\mathbb{C}^2$ can be identified with the closed unit ball in $\mathbb{R}^3$. The unit sphere in $\mathbb{R}^3$ is called the **Bloch sphere** in this context. It contains all extremal density matrices, i.e., all pure density matrices, i.e., all rank-one orthogonal projections. Any density matrix can be written as a convex combination of pure density matrices, and here we see that in the case of $\mathcal{H} = \mathbb{C}^2$, any density matrix can be written as a convex combination of (not more than) two pure density matrices.

III.2. Classical and Quantum Computation

Classical Computation. Now we turn away from physics but describe the framework of classical computation (by a computer) as if this was a physical system. The role of the particles is now played by *bits ($N = 1$)* or *bytes ($N \in \mathbb{N}$, $N \geq 2$)*, which we intepret as points $\underline{\sigma} = (\sigma_1, \ldots, \sigma_N) \in \Omega^{(N)}$ moving in the configuration space $\Omega^{(N)} := \{0,1\}^N$. A computation is a change of such a byte in time. Since computations are carried out in steps - not continuously, time is measured by integral numbers. That is, the state of a computation at time $t \in \mathbb{N}_0$ is

$$\underline{\sigma}(t) \;=\; \big(\sigma_1(t), \sigma_2(t), \ldots, \sigma_N(t)\big) \;\in\; \Omega^{(N)}\,. \tag{III.47}$$

Computations are trajectories $\underline{\sigma} : \mathbb{N}_0 \to \Omega^{(N)}$ of discrete time $t \in \mathbb{N}_0$, taking values in the finite set $\Omega^{(N)}$, $|\Omega^{(N)}| = 2^N$. The computation proceeds by applying a dynamical law to determine $\underline{\sigma}(t)$ from $\underline{\sigma}(t-1)$,

$$\forall\, t \in \mathbb{N}: \quad \underline{\sigma}(t) \;=\; F_t[\,\underline{\sigma}(t-1)]\,, \tag{III.48}$$

where $F_t : \Omega^{(N)} \to \Omega^{(N)}$, for any $t \in \mathbb{N}$. Any real-valued map $A : \Omega^{(N)} \to \mathbb{R}$ on the configuration space of all bytes defines an observable whose value at time $t \in \mathbb{N}_0$ for a given initial value $\underline{\sigma}(0) \in \Omega^{(N)}$ is given by

$$A_t[\underline{\sigma}(0)] \;:=\; A[\underline{\sigma}(t)] \;=\; A_t \circ F_t \circ F_{t-1} \circ \cdots \circ F_1[\underline{\sigma}(0)] \tag{III.49}$$

Probability in Classical Computations. Many problems in computation are naturally formulated in a probabilistic framework. One of these situations occurs in case our task is to

determine the minimum E_0 of a given function $H : \Omega^{(N)} \to \mathbb{R}$ and the set M of minimizers, i.e.,

$$M \;:=\; \big\{\underline{\sigma} \in \Omega^{(N)} \;\big|\; H(\underline{\sigma}) \;=\; E_0\big\}\,. \tag{III.50}$$

Among the methods to compute E_0 and M is *Simulated Annealing* or the *Monte Carlo Algorithm*, which we briefliy describe here.

(1) One chooses a starting point $\underline{\sigma}(0) = \big(\sigma_1(0), \ldots, \sigma_N(0)\big) \in \Omega^{(N)}$ and evaluates $H[\underline{\sigma}(0)]$. (The choice of the starting point may be random, but for many problems it is decisive to make a good guess which is not too far away from M.)

(2) For $t \in \mathbb{N}_0$ choose an index $j \in \mathbb{Z}_1^n$ randomly and set $\underline{\sigma}'(t+1) = \big(\sigma_1(t+1), \ldots, \sigma_N(t+1)\big) \in \Omega^{(N)}$ such that it differs from $\underline{\sigma}(t)$ exactly at the j^{th} position.

(3) Evaluate $H[\underline{\sigma}(t+1)]$.

(3a) If $H[\underline{\sigma}(t)] \geq H[\underline{\sigma}'(t+1)]$ then $\underline{\sigma}(t+1) := \underline{\sigma}'(t+1)$.

(3b) Conversely, if $H[\underline{\sigma}(t)] < H[\underline{\sigma}'(1)]$ then

$$\underline{\sigma}(t+1) \;:=\; \begin{cases} \underline{\sigma}'(t+1) & \text{with probability} \quad e^{-\beta\{H[\underline{\sigma}(t)]-H[\underline{\sigma}'(t+1)]\}}\,, \\ \underline{\sigma}(t) & \text{with probability} \quad 1 - e^{-\beta\{H[\underline{\sigma}(t)]-H[\underline{\sigma}'(t+1)]\}}\,. \end{cases} \tag{III.51}$$

Now replace t by $t+1$ and repeat the procedure from (2) on.

It can be proved that the trajectory $\big(\underline{\sigma}(t)\big)_{t\in\mathbb{N}_0}$ generated by the algorithm above concentrates on M. We do not go into detail here but only note that, while the framework is probabilistic, the computations carried out here are classical.

Quantum Computations. For quantum computers, we proceed in analogy to quantum mechanics: A configuration of the computer specified by a single bit $\sigma \in \Omega^{(1)} = \{0,1\}$ is replaced by a complex function $\psi(\sigma) \in \mathbb{C}$ of this bit. The configuration space $\Omega^{(1)} = \{0,1\}$ of the bit is hence replaced by the Hilbert space $\mathcal{H}^{(1)} := \ell^2(\Omega^{(1)}) \cong \mathbb{C}^2$ and the configuration $\psi \in \mathcal{H}^{(1)}$ is called **qubit**. Here, $V \cong W$ denotes isomorphy of Hilbert spaces.

Likewise, the configuration space of $N \in \mathbb{N}$ bits is replaced by the Hilbert space of N qubits,

$$\mathcal{H}^{(N)} \;:=\; \ell^2(\Omega^{(N)}) \;\cong\; \mathbb{C}^{2N} \;\cong\; \bigotimes^{N} \mathbb{C}^2\,, \tag{III.52}$$

The state of a quantum computer at time $t \in \mathbb{N}_0$ is described by a density matrix

$$\rho(t) \;\in\; \mathcal{DM}\big(\mathcal{H}^{(N)}\big)\,. \tag{III.53}$$

A quantum computation of $T \in \mathbb{N}$ steps is a family $u : \mathbb{Z}_1^T \to \mathcal{U}(\mathcal{H}^{(N)})$ such that $\rho(t)$ results from $\rho(t-1)$ by conjugation with $u(t) \in \mathcal{U}(\mathcal{H}^{(N)})$, for any time $t \in \mathbb{Z}_1^T$, that is

$$\rho(t) \;=\; u(t)\,\rho(t-1)\,u^*(t) \;=\; u(t)\cdots u(1)\,\rho(0)\,u^*(1)\cdots u^*(t)\,. \tag{III.54}$$

In particular, the final state is obtained from a unitary transformation of the initial state, too. More specifically,

$$\rho(T) \;=\; U(T)\,\rho(0)\,U^*(T)\,, \quad \text{with} \quad U(T) \;:=\; u(t)\cdots u(1) \;\in\; \mathcal{U}(\mathcal{H}^{(N)})\,. \tag{III.55}$$

Expectation values of states of quantum computers are defined just as in quantum mechanics: If $M = M^* \in \mathcal{B}(\mathcal{H}^{(N)})$ is a bounded self-adjoint operator representing an observable on N qubits, its expectation value in the state $\rho \in \mathcal{DM}(\mathcal{H}^{(N)})$ is given by

$$\langle M \rangle_\rho \;:=\; \mathrm{Tr}(\rho\, M)\,. \tag{III.56}$$

We also transfer the concept of measurement in quantum mechanics to quantum computation: We can access the state ρ only by expectation values (III.56) of observables $M = M^* \in \mathcal{B}(\mathcal{H}^{(N)})$. If a measurement is carried out, the state ρ is changed in such a way that it contains less information than before the measurements. We do not go into detail about this difficult conceptual problem but simply note that we can usually make only one single measurement with a given state. To obtain a reliable result one is bound to use redundancy, e.g., by preparing many identical copies of the initial state, run the quantum computation and make the same measurement many times, and eventually determine the correct result by using their statistics.