Chapter 2

(A Brief) Introduction to Elementary Particles

2.1 Standard Model Phenomenology

I shall here give a very brief overview of some important developments and discoveries in general physics and elementary particle physics relevant to the introduction to nuclear physics that will follow suite. Some of the material for the following section is taken from D. Griffiths, "Introduction to Elementary Particles".

2.1.1 Historical Notes

In 1932, the world of elementary particles was simple: The only known particles were the heavy proton (p) and neutron (n), the light electron (e), and the carrier of electromagnetic interactions, the photon (γ). What followed in the next three decades was a period of intense discovery of formerly unknown particles. The world of physics was swamped with discoveries in such a manner that Willis Lamb ('Lamb shift') joked in 1955: "For some time, a Nobel Prize was given out for the discovery of a new particle. Nowadays, someone who finds a new particle should be fined 10.000 $."

Until the 1950s, experimental particle physics was largely confined to the study of cosmic radiation and its atmospheric products. In 1952 the first modern particle accelerator went into activity, the 'Brookhaven Cosmotron' on Long Island, just outside New York. It could accelerate particles up to roughly 1 [GeV] kinetic energy (for comparison, the Large Hadron Collider currently reaches 6500 [GeV] kinetic energy per beam.)

2.1.2 Mesons

Mesons are "medium-weight" particles, hypothesized by H. Yukawa in 1934, that play a role in nuclear physics. To understand Yukawa's idea we take a look at a general parameterization of the four known forces of nature in terms of their range. Neglecting force constants, the following proportionality holds:

$$F_i \propto \frac{e^{-\frac{r}{a}}}{r^2} \tag{2.1}$$

where r is the distance between bodies and a is a range parameter that roughly takes on the following values:

i: electromagnetic, gravitational	$a \longrightarrow \infty$
i: strong	$a \approx 1$ [fermi] $= 10^{-15}$ [m]
i: weak	$a \approx 10^{-16} \ldots 10^{-17}$ [m]

So for electromagnetism and gravity the force follows a $\frac{1}{r^2}$ law. For the strong and the weak interactions it becomes very small when r becomes greater than characteristic nuclear/nucleon length scales (the proton radius is roughly $r_p \approx 0.88 \times 10^{-15}$ [m]).

Based on these experimental notions and basic quantum mechanical arguments, Yukawa proposed that there should be a particle that is exchanged between nucleons, responsible for the stability of an atomic nucleus. The nuclide ^{4}He ($2p$, $2n$), was known to be stable, despite the electromagnetic repulsion of the two protons. A "strong" nuclear force, surpassing the electromagnetic force in strength at nuclear length scale, had to be responsible for this. With the idea of the range of the interaction being inversely related to the rest mass of the force mediator, Yukawa argued that a quite heavy particle should be the mediator of the strong force.

Quantum theory demands that on a time scale Δt there is an uncertainty of measured energy according to

$$\Delta E \, \Delta t \geq \frac{\hbar}{2} \tag{2.2}$$

Since energy is related to mass, the rest mass of the mediator particle could be estimated from the time scale of its transmission. Using the proton radius r_p, a minimum value for this time scale is

$$\Delta t > \frac{r_p}{c} \approx 0.33 \times 10^{-23} \ \text{[s]} \tag{2.3}$$

With m_π the mass of Yukawa's meson, the energy fluctuation is given as $\Delta E = m_\pi c^2$, and so

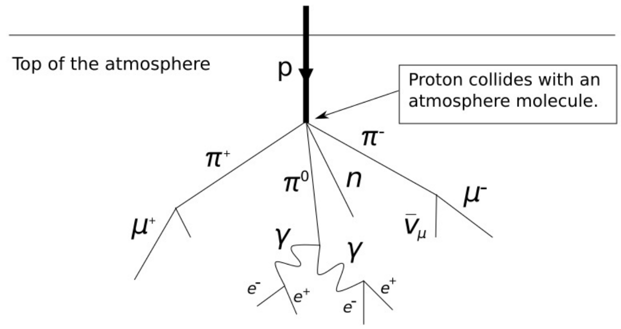
$$\begin{aligned} m_\pi c^2 \, \Delta t &> \frac{\hbar}{2} \\ \Rightarrow m_\pi &> \frac{10^{-34}}{2 \times 9 \times 0.33 \times 10^{16} \times 10^{-23}} \ \text{[S.I.]} \\ &\approx 0.185 \times 10^{-27} \ \text{[S.I.]} \end{aligned}$$

which translates into an upper bound for the rest mass of the π meson

$$m_\pi > 104 \left[\frac{\text{MeV}}{c^2}\right] \tag{2.4}$$

Today, the rest mass of the π mesons is known to be $\approx 135 \left[\frac{\text{MeV}}{c^2}\right]$ which shows that Yukawa's estimate was quite good. Comparing this rest mass with the rest masses of electron and proton, $m_e \approx 0.51 \left[\frac{\text{MeV}}{c^2}\right]$ and $m_p \approx 940 \left[\frac{\text{MeV}}{c^2}\right]$, the term "middle weight", or meson, is explained.

The π was detected in 1947 through cosmic radiation.



2.1.3 Antimatter - Dirac Equation

The most active period concerning the detection of antimatter particles was between 1930 and 1956. What does a course on 'relativity and nuclear physics' have to do with antimatter? Well, the existence of antimatter particles was **predicted** by **Dirac** as a consequence of his famous equation, and it happens to be the equation of motion of nucleons (as well as of all other massive fermions, like electrons). So in my mind, a 3rd-year course on nuclear physics cannot get around the Dirac equation.

Let us, however, begin with the developments that lead to this equation. The first steps were taken by Oskar Klein and Walter Gordon.

2.1.3.1 Klein-Gordon Theory

2.1.3.1.1 Derivation of the Klein-Gordon Equation

Starting point is the time-dependent Schrödinger equation (SEQ):

$$\imath\hbar \frac{\partial}{\partial t} \Psi(\mathbf{x}, t) = \hat{H}\, \Psi(\mathbf{x}, t) \tag{2.5}$$

is a differential equation first order in time and second order in space and, therefore, cannot be Lorentz covariant. To understand this consider the one-dimensional SEQ for a free particle, slightly rewritten:

$$\hbar \left(\imath \frac{\partial}{\partial t} + \frac{\hbar}{2m} \frac{\partial^2}{\partial x^2} \right) \Psi(x, t) = 0 \tag{2.6}$$

We know that $\partial^\mu\partial_\mu$ is a Lorentz scalar, so $a\,\frac{\partial}{\partial t} + b\,\frac{\partial^2}{\partial x^2}$ with constant a, b cannot be one.

As a first step, Klein and Gordon took the time derivative of Eq. (2.5)

$$\imath\hbar\,\frac{\partial^2}{\partial t^2}\,\Psi(\mathbf{x},t) = \hat{H}\,\frac{\partial}{\partial t}\,\Psi(\mathbf{x},t). \tag{2.7}$$

Note that we are working with a time-independent Hamiltonian operator in the Schrödinger picture. After introduction of Eq. (2.5) in the form $\frac{\partial}{\partial t}\,\Psi(\mathbf{x},t) = \frac{1}{\imath\hbar}\,\hat{H}\,\Psi(\mathbf{x},t)$ into the right-hand side of Eq. (2.7) we get

$$\imath\hbar\,\frac{\partial^2}{\partial t^2}\,\Psi(\mathbf{x},t) = \frac{1}{\imath\hbar}\,\hat{H}^2\,\Psi(\mathbf{x},t). \tag{2.8}$$

This is so far not a deviation from non-relativistic quantum physics, and Eq. (2.8) is still not homogeneous in time and space derivatives when considering the non-relativistic energy-momentum relation[1].

Now, Klein and Gordon argued that Einstein's energy of a particle in relativistic mechanics is $E = \pm\sqrt{\mathbf{p}^{\,2}c^2 + m_0^2c^4}$. With the usual prescription for the momentum operator in quantum theory, $\mathbf{p} \longrightarrow -\imath\hbar\,\hat{\boldsymbol{\nabla}}$ from which $\hat{\mathbf{p}}^2 = -\hbar^2\,\boldsymbol{\nabla}^2$. Making the corresponding replacement in the Hamiltonian in Eq. (2.8) and considering rest energy simply as a multiplicative constant,

$$\begin{aligned}\left(\hbar^2\,\frac{\partial^2}{\partial t^2} - \hbar^2c^2\,\boldsymbol{\nabla}^2 + m_0^2c^4\right)\Psi(\mathbf{x},t) &= 0\\ \left(\frac{1}{c^2}\,\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2 + \frac{m_0^2c^2}{\hbar^2}\right)\Psi(\mathbf{x},t) &= 0,\end{aligned} \tag{2.9}$$

which is equivalent to the procedure in non-relativistic quantum mechanics where the non-relativistic energy momentum relation is used at this point.

[1]One might ponder over using the square root directly in Eq. (2.5), but this results in operator roots and a mathematically quite complicated equation.

In Minkowski space with a metric tensor

$$\{g_{\mu\nu}\} = \{g^{\mu\nu}\} := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{2.10}$$

and real-valued coordinates, so $\{x^\mu\} := \begin{pmatrix} x_0 = ct \\ \mathbf{x} \end{pmatrix}$, the time derivative can be written like

$$\frac{1}{c}\frac{\partial}{\partial t} = \frac{1}{c}\frac{\partial x_0}{\partial t}\frac{\partial}{\partial x_0} = \frac{1}{c}\,c\,\frac{\partial}{\partial x_0} = \frac{\partial}{\partial x_0} =: \partial^0 \tag{2.11}$$

and thus Eq. (2.9) becomes

$$\begin{aligned} \left(\frac{\partial^2}{\partial x_0^2} - \sum_k \frac{\partial^2}{\partial x_k^2} + \frac{m_0^2 c^2}{\hbar^2}\right)\Psi(\mathbf{x}, t) &= 0 \\ \left(\partial^\mu \partial_\mu + \frac{m_0^2 c^2}{\hbar^2}\right)\Psi(x) &= 0 \end{aligned} \tag{2.12}$$

the **Klein-Gordon equation** where in the last step the position four-vector x replaces the coordinates $\mathbf{x}, t$ in the argument of the wave function. The operator $\Box = \partial^\mu \partial_\mu$ is called the **d'Alembertien**. The KG equation is manifestly Lorentz covariant, *i.e.*, it retains its form under Lorentz transformations. As can be shown straightforwardly, its solutions correspond to the correct relativistic energies of free particles of rest mass m_0. Note also that in the limit $m_0 \longrightarrow 0$ Eq. (2.12) yields the wave equation of electromagnetism.

2.1.3.1.2 Problems with Klein-Gordon Theory

It was, however, quickly realized that Klein and Gordon could not claim victory in having solved the problem of formulating the fundamental equation of relativistic quantum mechanics. The KG wave function

$\Psi_{\mathrm{KG}}(x)$ is a scalar field, but fermions have spin and a scalar wavefunction cannot describe two spin degrees of freedom, *i.e.*, spin projection "up" and spin projection "down". Perhaps particle spin could just be multiplied onto the wavefunction like in Schrödinger-Pauli theory. But this is certainly not satisfactory. The graver problem with Eq. (2.12) is, however, as follows.

In analogy to Schrödinger theory the conservation of the probability density is assured by it satisfying a continuity equation with an appropriate probability current density. For Klein-Gordon theory, this current density is

$$j^{\mu} := \frac{\imath\hbar}{2m_0}\left[\psi^*\left(\partial^{\mu}\psi\right) - \left(\partial^{\mu}\psi\right)^*\,\psi\right] \tag{2.13}$$

which shall be demonstrated.

Proof.

$$\partial_{\mu}j^{\mu} = \frac{\imath\hbar}{2m_0}\left[\left(\partial_{\mu}\psi^*\right)\left(\partial^{\mu}\psi\right) + \psi^*\left(\partial_{\mu}\partial^{\mu}\psi\right) - \left(\partial_{\mu}\partial^{\mu}\psi\right)^*\,\psi - \left(\partial^{\mu}\psi\right)^*\left(\partial_{\mu}\psi\right)\right] \tag{2.14}$$

The last term in Eq. (2.14) can be rewritten as

$$\begin{aligned} -\left(\partial^{\mu}\psi\right)^*\left(\partial_{\mu}\psi\right) &= -\left(\partial_{\nu}g^{\nu\mu}\psi\right)^*\left(\partial^{\kappa}g_{\kappa\mu}\,\psi\right) = -\left(\partial_{\nu}\psi\right)^* g^{\nu\mu}g_{\mu\kappa}\left(\partial^{\kappa}\psi\right) \\ &= -\left(\partial_{\nu}\psi\right)^*\delta^{\nu}_{\kappa}\left(\partial^{\kappa}\psi\right) = -\left(\partial_{\nu}\psi\right)^*\left(\partial^{\nu}\psi\right) \\ &= -\left(\partial_{\nu}\psi^*\right)\left(\partial^{\nu}\psi\right) \end{aligned} \tag{2.15}$$

and so cancels with the first term. Using the KG equation Eq. (2.12) in the second and third terms of Eq. (2.14) results in

$$\begin{aligned} \partial_{\mu}j^{\mu} &= \frac{\imath\hbar}{2m_0}\left[\psi^*\left(-\frac{m_0^2c^2}{\hbar^2}\,\psi\right) - \left(-\frac{m_0^2c^2}{\hbar^2}\,\psi^*\right)\psi\right] \\ &= \frac{\imath\hbar}{2m_0}\left[-\frac{m_0^2c^2}{\hbar^2} + \frac{m_0^2c^2}{\hbar^2}\right]\psi^*\psi = 0 \end{aligned} \tag{2.16}$$

which shows that the chosen probability current density in Eq. (2.13) indeed leads to a consistent KG theory. Since the probability current density four vector has the form given in Eq. (1.119) the KG probability density is the time-like component of Eq. (2.13)

$$\begin{aligned}\rho_{\mathrm{KG}} = \frac{j^0}{c} &= \frac{\imath\hbar}{2m_0c}\left[\psi^*\left(\partial^0\psi\right) - \left(\partial^0\psi\right)^*\psi\right] \\ &= \frac{\imath\hbar}{2m_0c}\left[\psi^*\left(\frac{1}{c}\frac{\partial}{\partial t}\psi\right) - \left(\frac{1}{c}\frac{\partial}{\partial t}\psi\right)^*\psi\right].\end{aligned} \tag{2.17}$$

The solutions of the KG equation are of plane-wave type and can be written as

$$\psi_{\mathrm{KG}}(x) = Ae^{\imath(\omega t-\mathbf{k}\cdot\mathbf{x})} + Be^{-\imath(\omega t-\mathbf{k}\cdot\mathbf{x})} \tag{2.18}$$

which when inserted into the KG equation yields the energy-momentum relation of relativistic theory. Using Eq. (2.18) in the representation of the KG density, Eq. (2.17), results in

$$\rho_{\mathrm{KG}} \propto -|A|^2 + |B|^2 \tag{2.19}$$

which is not hard to demonstrate. this means that the KG probability density could, depending on specific initial conditions which determine the coefficients A and B, become negative! This, however, is in stark contradiction with the fundamentals of quantum mechanics where probability density is always positive definite!

Historically, the KG equation was, therefore, regarded as a complete failure and discarded[2].

2.1.3.2 Dirac Theory

2.1.3.2.1 Derivation of the Dirac Equation

Dirac realized that the essential problem with the KG equation was the fact that it was a second-order differential equation. Thus, his idea was to formulate a Lorentz covariant **first-order** differential equation that treated space and time coordinates on an equal footing.

[2]As a little historical anecdote, a participant at one of the Solvay conferences in Brussels at the time is reported to have asked Dirac what he was currently working on. Dirac replied "I'm trying to construct a relativistic wave equation for the electron." Participant: "But Klein and Gordon already solved that problem!" They had not, and Dirac knew full well.

He started out by introducing a set of general parameters $\{\gamma^\mu\}$, the properties of which had to be determined, to see if the square root of the d'Alembertien could be developed in such a way that it takes on the form of a Lorentz scalar[3]. By postulate,

$$\begin{aligned}\Box = \partial_0\partial_0 - \sum_k \partial_k\partial_k \;&=:\; \gamma^\mu\gamma^\nu\,\partial_\mu\partial_\nu \\ &= \gamma^0\gamma^0\,\partial_0\partial_0 + \gamma^1\gamma^1\,\partial_1\partial_1 + \gamma^2\gamma^2\,\partial_2\partial_2 + \gamma^3\gamma^3\,\partial_3\partial_3 \\ &\quad + \left\{\gamma^0,\gamma^1\right\}\,\partial_0\partial_1 + \left\{\gamma^0,\gamma^2\right\}\,\partial_0\partial_2 + \left\{\gamma^0,\gamma^3\right\}\,\partial_0\partial_3 \\ &\quad + \left\{\gamma^1,\gamma^2\right\}\,\partial_1\partial_2 + \left\{\gamma^1,\gamma^3\right\}\,\partial_1\partial_3 + \left\{\gamma^2,\gamma^3\right\}\,\partial_2\partial_3\end{aligned}$$

where $\{,\}$ is the anti-commutator. This would be correct if

$$\{\gamma^\mu,\gamma^\nu\} = 0 \qquad \forall \mu \neq \nu \tag{2.20}$$

$$\left\{\gamma^0,\gamma^0\right\} = 2 \tag{2.21}$$

$$\left\{\gamma^k,\gamma^k\right\} = -2 \qquad \forall k \in \{1\ldots 3\} \tag{2.22}$$

For example, since $\left\{\gamma^0,\gamma^0\right\} = 2(\gamma^0)^2 = 2 \;\Rightarrow \gamma^0 = 1$ which is in accord with the above decomposition. The quantities γ are today known as **Clifford numbers** and the equations (2.20), (2.21), (2.22) as the conditions for a **Clifford algebra**, in this case the Clifford algebra of Dirac theory.

Let's take a look at the first condition, Eq. (2.20). If γ^κ is just a complex scalar, $\gamma^\kappa \in \mathbb{C}$, then $\{\gamma^\mu,\gamma^\nu\} = 2\gamma^\mu\gamma^\nu = 0$, since scalars commute. This can only be solved by setting $\gamma^\mu = 0 \;\; \forall\mu$ (or likewise $\gamma^\nu = 0 \;\; \forall\nu$) which does not lead to a valid set of parameters. Dirac concluded that the set $\{\gamma^\mu_{n\times n}\}$ had to be $n \times n$ matrices in order to arrive at anticommutators that yield zero!

Dirac's first try was to use the set $\{\mathbf{1},\boldsymbol{\sigma}\}$ which form a basis for all complex 2×2 matrices. However, neither this choice leads to a valid

[3] Remember that $\Box = \partial^\mu\partial_\mu$ and so $\sqrt{\Box} \neq \partial^\mu$ and $\neq \partial_\mu$. $\Box$ is a Lorentz scalar, and so is $\sqrt{\Box}$, but ∂^μ and ∂_μ are not Lorentz scalars!

decomposition of the d'Alembertien[4].

Dirac found the simplest set to be

$$\gamma^0 := \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix} \qquad \gamma^k := \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma}_k \\ -\boldsymbol{\sigma}_k & \mathbf{0} \end{pmatrix} \tag{2.23}$$

where the spin-Pauli matrices have been introduced[5]. Now it is straightforward to rewrite the d'Alembertien as

$$\left(\partial_0\partial_0 - \sum_k \partial_k\partial_k\right) \mathbb{1}_4 = \gamma^\mu \gamma^\nu \partial_\mu\partial_\nu = \gamma^\mu\partial_\mu \gamma^\nu\partial_\nu = \left(\gamma^\mu\partial_\mu\right)^2 \tag{2.24}$$

and so

$$\sqrt{\Box}\, \mathbb{1}_4 = \gamma^\mu\partial_\mu \tag{2.25}$$

Note that we here have obtained $\Box$ as the product of two (formally) Lorentz scalars in a four-dimensional vector space. This is no longer the same thing as the Minkowski space scalar product $\partial_\mu\partial^\mu$ between the derivative four-vectors.

A direct consequence of this astonishing finding is that the wave-function onto which such an operator acts cannot be a scalar function. It has to be a four-component vector-like quantity.

Dirac now formulated his famous equation. We first rewrite the KG equation as

$$\left(\hbar^2\partial^\mu\partial_\mu + m_0^2c^2\right) \Psi(x) = 0 \tag{2.26}$$

The square root of the first operator becomes in Dirac's formulation

$$\sqrt{\hbar^2\partial^\mu\partial_\mu}\, \mathbb{1}_4 = \hbar\gamma^\mu\partial_\mu \tag{2.27}$$

and so, with the four-momentum operator[6] $\hat{p}_\mu = \imath\hbar\partial_\mu$, we can formulate the **Dirac equation** for a free particle of rest mass m_0

[4]It can be shown in a general manner that only matrices with dimension multiples of 4 are possible solutions. So, for example, dimension 8 matrices can be constructed, but the resulting theory is identical to the dimension 4 theory in physical content.

[5]This set of matrices is called the "standard representation" of Dirac matrices. Other representations related through unitary transformations of the standard matrices are possible as well.

[6]Note that $\hat{p}^0 = \imath\hbar\frac{1}{c}\frac{\partial}{\partial t}$ and $p^0 = \frac{E}{c}$ which reproduces the correspondence principle from quantum mechanics for the energy operator $p^0c \longrightarrow \hat{p}^0c = \imath\hbar\frac{\partial}{\partial t}$. Likewise, $\hat{p}^k = -\imath\hbar\frac{\partial}{\partial x^k}$.

$$(-\imath\hbar\,\gamma^\mu\,\partial_\mu + m_0 c\,\mathbb{1}_4)\;\underline{\Psi}(x) = \underline{0} \qquad (2.28)$$

which is a matrix equation where the wavefunction becomes a **four-component spinor** that depends on the position four-vector x (which contains the time coordinate). We indeed have a first-order differential equation where space and time coordinates are treated on an equal footing, *i.e.*, in a four-vector.

Of course, we have to solve and interpret this equation and talk about what a spinor exactly is.

Using the standard representation the Dirac equation can be written out more explicitly as

$$\begin{aligned} \left\{-\imath\hbar\left[\gamma^0\partial_0 + \gamma^1\partial_1 + \gamma^2\partial_2 + \gamma^3\partial_3\right] + m_0 c\mathbb{1}_4\right\}\underline{\Psi}(x) &= \underline{0} \\ \left\{-\imath\hbar\left[\begin{pmatrix} \mathbb{1}_2 & 0_2 \\ 0_2 & -\mathbb{1}_2 \end{pmatrix}\frac{\partial}{\partial x_0} + \sum_{k=1}^{3}\begin{pmatrix} 0_2 & \boldsymbol{\sigma}_k \\ -\boldsymbol{\sigma}_k & 0_2 \end{pmatrix}\frac{\partial}{\partial x_k}\right] + m_0 c\begin{pmatrix} \mathbb{1}_2 & 0_2 \\ 0_2 & \mathbb{1}_2 \end{pmatrix}\right\}\begin{pmatrix} \underline{\Psi}^U(x) \\ \underline{\Psi}^L(x) \end{pmatrix} &= \underline{0}. \end{aligned} \qquad (2.29)$$

The second line follows from the realization that the γ matrices have a 2×2 block structure, and so the entire equation can be written in this so-called "bi-spinor" form. The associated 2-spinors are called "upper" ($\underline{\Psi}^U(x)$) and "lower" ($\underline{\Psi}^L(x)$) 2-spinors. Their significance will become clear when solutions of the free-particle Dirac equation are investigated.

Now, since in position-space representation $\mathbf{p} = -\imath\hbar\,\boldsymbol{\nabla}$ it is convenient to rewrite the term involving the spin-Pauli matrices using the scalar product between the 3-vector of the Pauli matrices and the 3-vector of momentum, $\boldsymbol{\sigma}\cdot\mathbf{p} = \boldsymbol{\sigma}_x\hat{p}_x + \boldsymbol{\sigma}_y\hat{p}_y + \boldsymbol{\sigma}_z\hat{p}_z$,[7] and the Dirac equation becomes

[7]Note, e.g., that the product of a Pauli matrix with a scalar momentum operator is well defined: $\boldsymbol{\sigma}_x\hat{p}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\hat{p}_x = \begin{pmatrix} 0 & \hat{p}_x \\ \hat{p}_x & 0 \end{pmatrix}$.