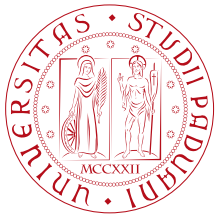

UNIVERSITÀ DEGLI STUDI DI PADOVA

FACOLTÀ DI SCIENZE MM.FF.NN.

DIPARTIMENTO DI MATEMATICA PURA ED APPLICATA

TESI DI DOTTORATO IN MATEMATICA

XXIII CICLO

PHYSICAL PROPERTIES OF CARBON NANOTUBES BY COMPUTATIONAL METHODS

Advisor: Prof. Maria Morandi Cecchi

Candidato: Vittorio Rispoli

Anno Accademico 2010/2011

Volli, e volli sempre, e fortissimamente volli

Riassunto

Il lavoro di ricerca della tesi di Dottorato ha avuto come oggetto la simulazione, dal punto di vista computazionale, delle proprietà fisiche dei Nanotubi di Carbonio (CNTs). Tale lavoro vuole essere un contributo alla viva e ricchissima ricerca, sia teorica che sperimentale, a livello mondiale riguardante le nanotecnologie e le loro incredibili applicazioni.

Per prime sono state studiate le proprietà meccaniche del nanotubi. In contesto di Dinamica Molecolare è stato possibile calcolare alcuni dei principali parametri strutturali caratterizzanti i CNT. Per la simulazione del comportamento meccanico di un singolo nanotubo, abbiamo considerato un modello continuo per ottenere una accurata descrizione dei potenziali inter-atomici rappresentanti le interazioni tra gli atomi del nanotubo soggetto ad un carico applicato su di esso. Il metodo più efficiente per la risoluzione numerica del problema è stato l' "Atomic Scale Finite Element Method", grazie al quale è stato possibile calcolare i valori del modulo di Young, del rapporto d Poisson, del modulo di rigidità ed altre importanti caratteristiche strutturali dei nanotubi di carbonio.

Il secondo argomento di cui ci siamo occupati sono state le proprietà elettriche dei nanotubi. Abbiamo approfondito i meccanismi di generazione della corrente all'interno di un nanotubo metallico e abbiamo simulato numericamente la dinamica degli elettroni per calcolare i valori della corrente al variare del potenziale applicato. Per simulare il comportamento elettrico dei nanotubi è stato necessario studiare la dinamica accoppiata di elettroni e fononi, questi ultimi generati durante l'evoluzione temporale del sistema. L'introduzione dei fononi nel modello è servita per catturare alcuni fenomeni quantistici presenti nel sistema studiato, tra cui, ad esempio, i processi di scattering elettrone-fonone, che influenzano in maniera significativa il moto degli elettroni e, in ultima analisi, i valori calcolati della corrente.

Sono stati studiati nanotubi di tipo metallico poiché questo, grazie anche all'alto livello di simmetria dei nanotubi, ha permesso di poter affrontare un problema con uno spazio delle fasi abbastanza ridotto. La dinamica accoppiata di elettroni e fononi è stata simulata tramite equazioni cinetiche di trasporto di tipo Boltzmann, con le quali abbiamo descritto l'evoluzione temporale delle distribuzione delle particelle.

Il nostro contributo maggiore al modello fisico studiato è stato quello di introdurre una formula esplicita per calcolare, in maniera consistente con la dinamica del sistema durante tutta l'evoluzione temporale, i coefficienti di accoppiamento elettrone-fonone (EPC). In tutti i modelli presenti in letteratura, infatti, i parametri di EPC sono sempre considerati come costanti del problema, ottenuti o come parametri di fitting o tramite relazioni empiriche con altri dati per problema, per esempio dipendenti dal diametro del nanotubo. Nel nostro modello abbiamo invece stabilito una diretta dipendenza tra le lunghezze di scattering (che si usano nel calcolo dei valori di EPC) e le densità di distribuzione dei fononi, che ha consentito

di calcolare in ogni istante in maniera consistente con lo stato del problema i coefficienti necessari.

Tale approssimazione si è rilevata di grande importanza poiché durante l'evoluzione temporale del sistema le distribuzioni dei diversi modi di fononi presi in considerazione assumono valori in intervalli molto ampi e quindi non è corretto ipotizzare valori costanti per le lunghezze di scattering tra le particelle. I risultati ottenuti mostrano ottimo accordo confrontati con i dati sperimentali a disposizione in letteratura; le simulazioni basate sul modello migliorato hanno mostrato, inoltre, maggiore efficienza computazionale, soprattutto in termini di tempi di esecuzione.

I nanotubi di carbonio con diametro piccolo (al più $3\,nm$), cioè quelli considerati nel nostro modello, sono usati in un gran numero di applicazioni nanotecnologiche. Al fine di ottenere simulazioni utili anche per altre applicazioni pratiche, dovrebbero essere considerati anche nanotubi con diametro grande. I calcoli ottenuti estendendo in maniera automatica il modello studiato (che ha una dimostrazione di validità solo per nanotubi di raggio piccolo) non hanno dato risultati soddisfacenti, in alcuni casi molto lontani dai dati sperimentali. Secondo noi, quindi, ci sarà bisogno di cambiare il modello fisico usato per approssimare il problema; questo sarà un problema molto interessante come futuro argomento di ricerca.

Dal punto di vista matematico, il problema presentato è costituito da un sistema di Leggi di Conservazione iperboliche, multi-dimensionale e con termini di sorgente a membro destro. Il trattamento numerico di un problema di questo tipo è un compito piuttosto difficile, data la mancanza di risultati teorici generali che possano garantire l'accuratezza e l'affidabilità dei risultati ottenuti.

Applicando lo schema numerico usato per l'approssimazione numerica del problema differenziale iniziale, abbiamo trovato ottimi risultati anche in una serie di altri casi in contesti più generali. Anche se criteri qualitativi devono essere ancora dimostrati, il metodo proposto può sicuramente essere utilizzato come schema di riferimento affidabile per calcoli riguardanti problemi differenziali di tipo iperbolico.

Abstract

The subject of this PhD thesis is the simulation of Carbon Nanotubes (CNTs) physical properties by computational methods. The work presented in this thesis is a contribution to the world wide scientific research, both theorical and experimental, regarding nanotechnology and its incredible applications.

Our first topic of interest was the study of CNTs' mechanical properties. We considered a Molecular Dynamics setting for our computations, thanks to which we were able to compute most of the characterizing structural parameters of carbon nanotubes. The continuous model describing the physics of the system was given by accurate definitions of the inter-atomic potentials characterizing the interactions between nanotube's carbon atoms subject to the applied loads. To compute simulation results, the "Atomic Scale Finite Element Method" revealed to be very efficient. We were thus able to predict the values of the Young modulus, of the Poisson ratio, of the shear modulus and also of other important structural parameters of CNTs.

The second subject of interest were CNTs' electrical properties. Our aim was to simulate the current generation and compute current values inside a nanotube at whose ends an electric potential difference was applied. In order to model such behavior, it was necessary to study the coupled dynamics of electrons and phonons, the latter being generated during the time evolution of the system. Phonons were introduced in order to take into account quantum mechanics phenomena, such as the scattering between electrons and phonons, that strongly influence the behavior of the electrons and, thus, the computed current values.

We considered only metallic nanotubes; for this reason and thanks to CNTs high symmetry, it was possible to restrict to a low-dimensional phase-space problem. The dynamics of electron and phonons was modeled by kinetic Boltzmann-type transport equations, governing the time evolution of particles distributions.

Our main contribution to the physical modeling of this problem was to introduce an explicit formula to compute in a self-consistent way Electrons-Phonons Coupling (EPC) parameters during the time evolution of the system. In all models available in the literature, indeed, EPC coefficients were taken as fixed constant values, obtained as fitting parameters or by simple scaling with the diameter. We defined, instead, a direct dependence of the scattering lengths (which are used for the computation of the EPC values) from phonons densities, thus computing coupling parameters self-consistently depending on system states.

This was a fundamental step since during the time evolution of the system, the densities of all the different considered phonon modes assume very large ranges of values and, for this reason, it is not correct to assume constant values of the coupling coefficients. The computed results were in very good agreement with experimental data and simulations based on the

improved model also showed a significantly decreased computational cost to compute the desired solution.

Carbon nanotubes just a few nanometers in diameter, which are those considered in our model, are used in many nanotechnological applications. To find simulation results that could be used for other practical applications, large diameter nanotubes should also be considered. Computations obtained extending the actual model (which is proved to be correct only for small diameter tubes) in a straightforward way are not accurate enough and thus a different physical model needs to be considered. This will be a very interesting subject for future investigations.

From the mathematical point of view, the given problem constitutes a system of multi-dimensional hyperbolic Conservation Laws with source terms at the right hand side. The numerical treatment of a problem of this type is a difficult task, given the lack of a general theory which could guarantee reliability of the computed results.

Thanks to the presented scheme for the numerical approximation of the differential problem, we found very good results also in a wide range of more general situations. Even if qualitative criteria need yet to be proved, the proposed method can be used as a reliable reference scheme for computations of this type.

Contents

Introduction

Nanotechnology is the forefront of modern technology. A great number of nanotechnological products have already been produced thanks to great advances in our ability to manipulate such small objects; a comprehensive and up-to-date list of all the already available products is publicly accessible on-line at www.nanotechproject.org/inventories/consumer. The ability to modify objects down to the atomic level opens a whole new world of possibilities and applications not even conceivable before the nano era.

It is very hard to define what nanotechnology is; even if the term nanotechnology is on stage since the 1959 famous speech of Feynman ([**2**]) and the first "nano-based" application is already as old as almost 20 years old, still nowadays there is not, among the scientific community, complete agreement on the definition of this term. One of the key point of the argue is to decide what *should* be considered a nanotechnological device and what *should not*; this is a very tricky question since anything is made of atoms and any property an object can have, ultimately depend on its atomic structure. The most important requirement for nanotechnology definition is that the nano-structure has special properties that are exclusively due to its nanoscale proportions. So, what exactly is nanotechnology? A good definition could be ([**10**]):

"The design, characterization, production, and application of structures, devices, and systems by controlled manipulation of size and shape at the nano-meter scale (atomic, molecular, and macromolecular scale) that produces structures, devices, and systems with at least one novel/superior characteristic or property."

All new objects, having some kind of manipulation at the nano-meter level in their production process and, thus, obtaining novel properties at the macroscopic level, can certainly be considered nanotechnological products. Just to name a few of the fields in which nanotechnology innovations are soon arriving: informatics (PCs hundreds of times faster then today), environmental sciences (automatic cleanup of existing pollution), material science (big magneto-resistance in nanocrystalline materials, nanoparticle reinforced materials), medicine (advanced drug delivery systems, biomedical prosthesis), new generation of lasers, and many others.

Nanotechnology is, mostly, based on nanosized components, objects whose dimensions (or at least one of them) range among the nanometric scale (recall that $1\ nm = 10^{-9}m$). Nowadays, many kinds of these components exist; the first one to be discovered by the group of Smalley, in 1985, was *Buckminsterfullerene*, a hollow sphere whose surface is a tiling of hexagons and pentagons with carbon atoms at every vertex.

Other fundamental components in nanotechnology are *Carbon Nanotubes* (CNTs therein), tubes made of Carbon atoms arranged on the surface of the tube in a hexagonal lattice, whose diameter range from tenths to tens of nanometer and that are from hundreds up to thousands nanometers in length.

CNTs have incredible physical properties, both mechanical and electrical, and they are the basis of most of the new nanotechnological innovations. When CNTs amazing physical properties were theoretically predicted, very soon a world wide interest rose up and both academic and industrial research started in this new field.

There is a long aged argue about who was the first to discover carbon nanotubes. It is commonly reported that their father is Sumio Iijima who described CNTs in 1991 ([**34**]). Nevertheless, scientific publications in which authors described small *graphitic tubules* were available long time before Iijima's work: just to name a few, Radushkevich and collaborators in Russia reported about carbon nanotubes as early as in 1952 ([**51**]) and Endo from Japan, in collaboration with Oberlin in France, reported on the observation of carbon nanotubes by electron microscopy in 1976 ([**46**]). Probably, Iijima's fortune is due to the fact that he was the first to be able to produce them; for the scientific community, to "give birth" to something is as important (if not even more important, indeed) as the discovery of it.

A similar history unite carbon nanotubes and *graphene*: it is a planar lattice of hexagons whose vertices consist in carbon atoms; it can be thought of as the bi-dimensional counterpart of graphite. Figure 1 shows, from the upper left corner and then clockwise: a graphene sheet, the three dimensional graphite, a Buckminsterfullerene and a Carbon Nanotube.

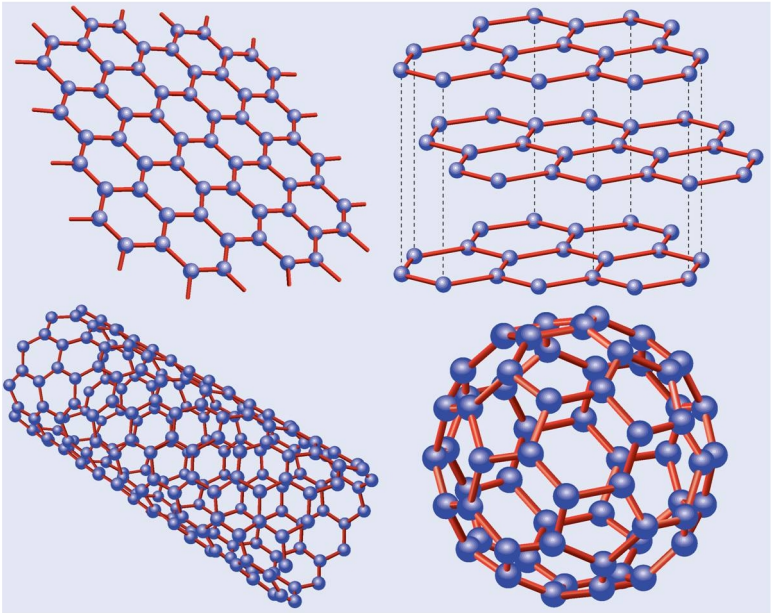

FIGURE 1. Picture from http://physics.bu.edu/ neto/mother.jpg

Single-layer carbon foils were described already in 1962 ([**12**]) but only in 2004 it was possible to create the first stable graphene sheet ([**25**]). The huge amount of theoretical work could, from that moment, be tested and a strong impulse was given to all related researches. The great importance of this discovery can be inferred from the fact that the authors were honored with the 2010 Nobel Price in Physics; as it can be read in the motivation, such

result came *"... at a time when many believed it was impossible for such thin crystalline materials to be stable"* ([**1**]).

Our work finds its context in this framework. We studied the physical, both mechanical and electrical, properties of CNTs. Since their discovery, CNTs have attracted great interest and huge work has been made to understand these new, fascinating as much as mysterious objects.

First of all, it took a very long time before it was possible to produce them in well-defined ways, i.e. isolated tubes or small blocks of aligned tubes, and still nowadays production techniques are a very active field of research. First theoretical works reporting on nanotubes physical properties date back to the beginning of the 90's and still today also this is a very rich field. Many difficulties have to be handled to obtain results from the very complex models regarding nanotubes: physical phenomena at the atomic scale are affected by quantum mechanical effects which have to be included in the models for CNTs; moreover, the scientific community had to wait a long time before experimental data were available, slowing down progresses very much. During the past ten years, instead, the knowledge and understanding of CNTs has had a great impulse, thanks to the great number of always more accurate data available from experiments.

We tried to give contributions to the understanding of CNTs properties, both mechanical and electrical, together with analyses and improvements of the numerical methods used in the mathematical modeling of CNTs physics.

Our first argument of interest was the characterization of CNTs mechanical properties. Because of their peculiar dimensions, nanotubes are considered as prototypes of mono dimensional systems, hence real objects which can be used to develop 1D physics theories. Moreover, again because of their size, nanotubes live at the frontier which separates continuum and molecular models. Physical models regarding mechanical properties of CNTs have been proposed in both Continuum Mechanics (CM) and Molecular Dynamic (MD) frameworks. In our study we adopted a quite new technique, called Atomic Scale Finite Element Method; it was presented in [**43**] and it is a MD based method. One of the main advantages of such method is that it comes from the same setting as standard Finite Element Methods and so it is possible to create very powerful hybrid methods. Dealing with an atomistic method, we found accurate and efficient representation of the 3D structure of a CNT: once reliable physical potentials, describing atomic interactions, were found we were able to compute some of the characterizing structural parameters of CNTs, such as Young's modulus and Poisson ratio. We could also describe the dependence of these properties on the different types of nanotubes, obtained varying the tube's spatial configuration, i.e. their radius or the distribution of the atoms on the surface of the tube ([**17**]).

The second part of our work was dedicated to the study of the electrical properties of carbon nanotubes. Their remarkable electrical properties make them, in many cases, the best candidates for innovative electronic applications; they can be used as charge storage components (i.e. in supercapacitors), as semiconductors, metals or even superconductors.

The first models regarding CNTs electric behavior were descriptions at a macroscopic level; it was soon clear that deeper models were needed to take into account, among other things, the significant quantum mechanics effects which deeply influence the behavior of electrons in conducting nanotubes. To model the behavior of the electrons in a nanotube subject to an applied bias, kinetic Boltzmann-type equations have been defined, governing

the time evolution of electrons' distribution inside the tube. We defined an improved model, improved in the sense that it described more accurately the real physical problem, assuming more accurate descriptions of the interactions between all the characters on the stage ([**16**]). Starting from such models, we could find numerical solutions which reproduced experimental data in a very accurate way.

To numerically simulate the evolution of the governing PDEs we had to face a system of hyperbolic Conservation Laws with source terms at the right hand side (conservation laws with source terms are usually called Balance Laws); in this already difficult setting, things were made harder by the multi-dimensionality of the problem. Even if a general theorical result could not be proved, our work shows the adopted scheme is efficient and reliable (i.e. stable and convergent) in a wide range of situations, thus being a valid reference for this kind of computations and a solid basis for the seek of a general theorical result.

The thesis is organized as follows: in Chapter 1, we first present the geometrical and structural characterization of CNTs and after the description of the AFEM scheme we conclude showing the obtained results about the mechanical properties of carbon nanotubes; in Chapter 2, we describe the physical setting in which the Boltzmann equations governing the dynamics of electrons are derived, together with a description of the quantum mechanics phenomena occurring in the time evolution of the system. In Chapter 3 we recall the general theory regarding the mathematical treatment of both hyperbolic conservation and balance laws and in Chapter 4 we recall the general theory on the numerical approximation of such type of equations. Finally, in the last chapter we present simulation results regarding electrons dynamics and, even if without a full theorical proof, numerical solutions of such a difficult hyperbolic Balance Laws problem, as that in the presented model, are given.

CHAPTER 1

Mechanical properties of Single Wall Carbon Nanotubes

Because of its simple atomic configuration, graphene physical properties are easier to study and analyze with respect to those of nanotubes. Nevertheless, thanks to the great similarity between the two structures it is often possible to use the knowledge about one of the two to state predictions about the other. This was the case, for example, for the structural properties of carbon nanotubes.

1. Classification

Carbon nanotubes can be found in nature as single standing tubes, in which case they are called *Single Wall Carbon Nanotubes* (SWCNTs), or as many tubules nested one into the other, called *Multi Wall CNTs* (MWCNTs). The (mean) distance among tubes in MWCNTs is similar to that of graphene planes which compose the three dimensional graphite.

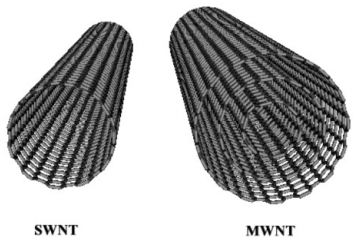

FIGURE 1. Single and Multi Wall Carbon Nanotubes

The geometrical description and the characterization of the structural properties of a Single Wall CNT is made easier starting from that of graphene ([**55**]). A SWCNT is usually described as a rolled-up graphene sheet; see Figure 2.

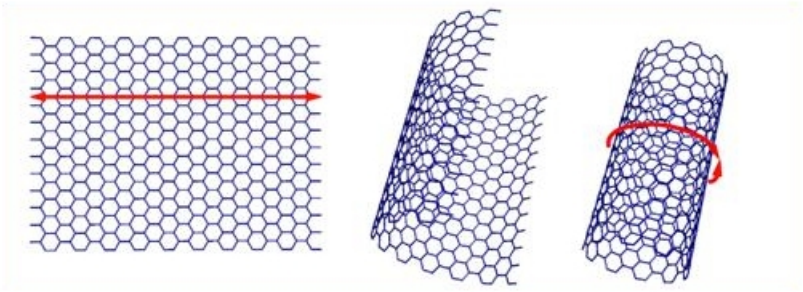

FIGURE 2. Rolling of a graphene sheet into a CNT

There are infinitely many ways to fold the honeycomb lattice of hexagons into a tube and most of them define different nanotubes. A vector connecting two *cristallographically* equivalent atoms (which are two atoms that will coincide once the plane is folded) correspond to a circumferential vector of the tube: such vector is called the *chirality vector* C_h.

A reference system can be fixed on the plane, whose origin coincides with one atom and with the two axis as in Figure 3.

One can then define two *lattice basis vectors* a_1 and a_2

$$a_1 = \left(\frac{3}{2}, \frac{\sqrt{3}}{2}\right) a_{CC}, \qquad a_2 = \left(\frac{3}{2}, -\frac{\sqrt{3}}{2}\right) a_{CC},$$

where a_{CC} is the (mean) Carbon-Carbon bond length

$$a_{CC} \approx 1.42\ \mathring{A} = 0.142\ nm$$

and a shift vector

$$s = (1, 0)\, a_{CC}.$$

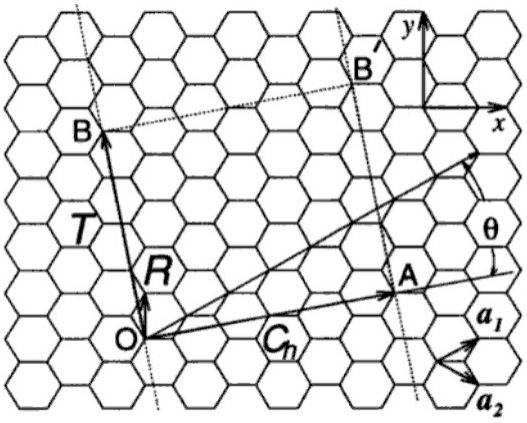

FIGURE 3. Lattice basis and Chirality vector C_h

Any atom on the lattice could be used to identify a tube. Anyway, to save the hexagonal "motif" on the tube, we have to consider only those atoms which are represented in the lattice basis reference system by integer components. Any CNT, in this way, can be represented by its chirality vector

$$C_h = n\, a_1 + m\, a_2 \equiv (n, m)$$

and if we further assume $0 \leq m \leq n$, we obtain a one to one correspondence. According to their chirality, SWCNTs are named: *"zigzag"* when they are identified by the couple $(n, 0)$, *"armchair"* when identified by the couple (n, n) and general *"chiral"* in all other cases.

Using the chirality vector, we can characterize all the geometrical and symmetrical properties of a nanotube. For example, we can easily compute the *diameter* of the tube

$$d_h = \frac{|C_h|}{\pi} = \frac{a_0}{\pi}\sqrt{n^2 + nm + m^2},$$

where $a_0 = \sqrt{3}a_{CC}$ is the lattice constant. It is also possible to compute the *chiral angle* θ:

$$\cos(\theta) = \frac{C_h \cdot a_1}{|C_h| \cdot |a_1|} = \frac{2n+m}{2\sqrt{n^2+nm+m^2}}$$

which is the angle between C_h and a_1; given the restriction $0 \leq m \leq n$, we have that $0 \leq \theta \leq \pi/6$. In particular, $\theta = 0$ for zigzag tubes and $\theta = \pi/6$ for armchair tubes.

We define the *translational* vector T_h as the shortest lattice vector perpendicular to the chiral vector (the one, say, closest to the y-axis positive direction); it can be easily computed by:

$$T_h = t_1\, a_1 + t_2\, a_2 \equiv (t_1, t_2),$$

where the coefficients t_1 and t_2 are integer numbers obtained imposing $C_h \cdot T_h = 0$. This leads to

$$t_1\,(2n+m) + t_2\,(n+2m) = 0,$$

which, since t_1 and t_2 are coprime, in turn gives

$$t_1 = \frac{n+2m}{d_R}, \qquad t_2 = -\frac{2n+m}{d_R},$$

where $d_R = \gcd(n+2m, 2n+m)$ is the greatest common divisor of the two numbers. Moreover,

$$|T_h| = \frac{\sqrt{3}\,|C_h|}{d_R}.$$

Vectors C_h and T_h define the *unit cell* of the nanotube, which is, in general, the smallest building block of a crystal, whose geometric arrangement defines a crystal's characteristic symmetry and whose repetition in space reproduces the crystal lattice. We can compute the area of the unit cell of a SWCNT by

$$A_S = |C_h \times T_h| = |C_h| \cdot |T_h| = \frac{\sqrt{3}\,a_0^2}{d_R}\left(n^2+nm+m^2\right)$$

while the area of a single hexagon (which is also the area of a unit cell of graphene) is given by the formula

$$A_G = |a_1 \times a_2| = \frac{\sqrt{3}\,a_0^2}{2}.$$

Looking at the hexagonal lattice on graphene, we can see that to count the number of atoms contained in a portion of the sheet we can count the number of cells in the considered area and multiply it by 2, since for each cell we can count two atoms. The same reasoning holds for a SWCNT and so we have that the total number of atoms inside the unit cell is given by:

$$N_C = 2\frac{A_S}{A_G} = 2\frac{(n^2+nm+m^2)}{d_R}.$$

The total number of atoms in the unit cell is a very important parameter for calculations regarding mechanical and electrical properties of SWCNTs and it can always be computed using this simple formula.

We recall in Table 1 the structural parameters described up to now.

Symbol	Name	Formula	Value
a_0	lattice constant	$a_0 = \sqrt{3}a_{CC} \approx 2.46\mathring{A}$	$a_{CC} = 1.42\mathring{A}$
$a_1,\ a_2$	basis vectors	$(\sqrt{3}/2, 1/2)a_0, (\sqrt{3}/2, -1/2)a_0$	
C_h	chiral vector	$C_h = n\,a_1 + m\,a_2 \equiv (n, m)$	$(0 \le m \le n)$
d_h	tube diameter	$d_h = \lvert C_h\rvert/\pi$	
θ	chiral angle	$\cos(\theta) = \langle C_h, a_1\rangle$	$0 \le \theta \le \frac{\pi}{6}$
T_h	translational vector	$T_h = t_1\,a_1 + t_2\,a_2$	
		$t_1 = (n + 2m)/g_R$	$t_2 = -(2n + m)/g_R$
N_C	unit cell C atoms	$N_C = 2(n^2 + nm + m^2)/g_R$	

TABLE 1. Structural parameters of CNTs

The high symmetry of atoms distribution in CNTs is heavily exploited in all predictions regarding their properties; depending on their chirality, carbon nanotubes have different symmetries and hence different *space groups*. This is an important classification because once it is possible to predict a property of a specific tube, it is then possible to extend it to all tubes in the same symmetry class ([**23, 24**]).

Exploiting the high symmetry of atoms picture on the surface of CNTs, it is possible to implement very efficient computational codes both for tubes spatial representations or for coordinate definitions in molecular dynamics computations ([**53**]).

2. Mechanical properties

We now present an atomistic computational method for simulations regarding CNTs mechanical properties, following the approach proposed in [**14**]. The developed method is based on the *Atomic-Scale Finite Element Method* (AFEM) recently proposed in [**43**] as an efficient alternative to molecular mechanics for the nonlinear minimization of the system energy, having comparable accuracy. This is due to the use of both first and second derivatives of system energy which improves the convergence of the method allowing the reduction of the total computational time. AFEM provides a reliable framework for efficient and accurate computation of mechanical properties of CNTs ([**43**]).

An advantage of AFEM is that it has the same formal structure of the continuum *Finite Element Method* (FEM), permitting the combination of AFEM with FEM when needed, avoiding artificial interfaces. For example, hybrid methods are applied in multiscale simulation for CNTs where FEM is used for the macro-scale problem while AFEM is used for the micro-scale problem ([**50**]). The idea is to combine the atomistic method where it is necessary to capture the nanoscale physical laws with the continuum FEM where it is possible, instead, to collect the behaviour of atoms reducing the degrees of freedom of the system.

2.1. Numerical modeling. In this section the formal modeling framework is presented: this includes the geometrical aspects, the specific molecular potential and the atomic-scale finite element method for carbon nanotubes.

2.1.1. *Lattice modeling.* Reflecting the repetitive display of the lattice, a numbering scheme can be defined to model CNTs with reference to the unstressed configuration of a graphene sheet, obtaining a description of the entire sheet as the tiling repetition of an elementary Y-shaped cell ([**14**]).

A sheet of graphene is defined as $\mathcal{G} = (\mathcal{A}, \mathcal{B}, \mathcal{C})$ where $\mathcal{A}$ is the set of all the atoms of the sheet, $\mathcal{B}$ is the set of all the binary bonds between pairs of adjacent atoms and $\mathcal{C}$ is the set of all the ordered couples of adjacent bonds.

Every atom $a \in \mathcal{A}$ is considered as a material point with two attributes: a label, lab_a, and its position at time t, $\text{pos}_a(t)$.

Because of the hexagonal structure of the graphene sheet $\mathcal{G}$ another useful coordinate system that is not orthonormal can be specified for the physical space: $j_1 = a_0(1, 0, 0)$, $j_2 = a_0\left(1/2, \sqrt{3}/2, 0\right)$, $j_3 = a_{(C-C)}(0, 0, 1)$. A rhomboidal grid on the graphene sheet is defined as $\mathcal{A}_1 \subset \mathcal{A}$:

$$\mathcal{A}_1 = \{a \in \mathcal{A} | \, \text{pos}_a(t) = \alpha_1 j_1 + \alpha_2 j_2, \, \alpha_1, \alpha_2 \in \mathbb{Z}\} \, .$$

Every carbon atom $a \in \mathcal{A}_1$ will be uniquely identified by the integer couple $(\alpha_1, \alpha_2) \in \mathbb{Z}^2$ and will be referred to as $\text{lab}_a = \mathcal{A}_1(\alpha_1, \alpha_2) = (\alpha_1, \alpha_2, 1)$. Let $\mathcal{A}_2 = \mathcal{A} \backslash \mathcal{A}_1$ be another rhomboidal grid with the same integer constraints. More precisely, $\mathcal{A}_2$ is the set whose atoms have their positions that are the shifted version of the atoms in $\mathcal{A}_1$. This shift is induced by the shifting vector $s = a_{(C-C)}\left(\sqrt{3}/2, 1/2, 0\right)$. Hence, $\mathcal{A}_2 = \{a \in \mathcal{A} | \, \text{pos}_a(t) = \alpha_1 j_1 + \alpha_2 j_2 + s, \, \alpha_1, \alpha_2 \in \mathbb{Z}\}$. The label for the atom $a \in \mathcal{A}_2$ is given by $\text{lab}_a = \mathcal{A}_2(\alpha_1, \alpha_2) = (\alpha_1, \alpha_2, 2)$.

The Y-shaped cell is the elementary tile that covers the entire graphene sheet ([**15**]). Every cell is uniquely identified by the ordered pair of integers $(\alpha_1, \alpha_2) \in \mathbb{Z}^2$ and consists of two atoms and three bonds. Every bond can be specified by an ordered pair of neighbouring atoms and, vice versa, one can refer to the first atom of the bond $b_u(\alpha_1, \alpha_2)$ as $b_u^1(\alpha_1, \alpha_2)$ for $u = 1, 2, 3$ and similarly for the second atom.

The length vector $B_u(t)$ can be associated with every bond $b_u(\alpha_1, \alpha_2)$ at time t as $B_u(t) = \text{pos}_{a_1}(t) - \text{pos}_{a_2}(t)$ with length $l_u(t) = |B_u(t)|$.

In addition to atoms and bonds, the presented numbering scheme is extended to include couples of bonds that share a common atom. Six couples of bonds are associated to every Y-shaped elementary cell (Figure 4): let a generic cell be $Y(\alpha_1, \alpha_2)$, then these couples are named $c_i(\alpha_1, \alpha_2)$ with $i = 1, \ldots, 6$.

Also in this case, one can refer to the first bond of the couple $c_u(\alpha_1, \alpha_2)$ as $c_u^1(\alpha_1, \alpha_2)$ for $i = 1, \ldots, 6$ and similarly for the second bond. Let $c = (b_1, b_2) \in \mathcal{C}$ be a couple of bounds that share a common atom, then the angle between the bonds is given by
$\rho_c(t) = (B_1^T(t) \cdot B_2(t))/(l_1(t) l_2(t))$, where B_i and l_i are, respectively, the vector and the length associated to the bond b_i for $i = 1, 2$. This numbering scheme applies also to SWCNTs ([**55**]).

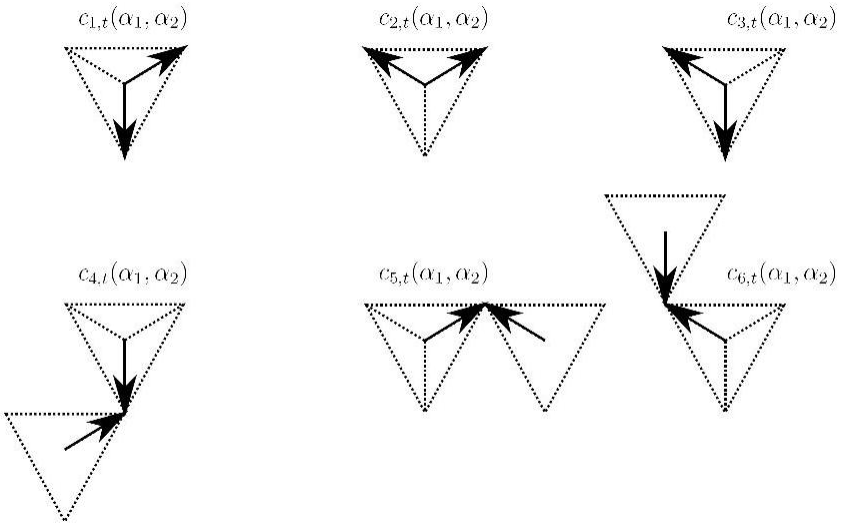

FIGURE 4. The six couples of angles c_i for $i = 1, \dots, 6$.

2.1.2. *Empiric potential.* The behaviour and the energetics of molecules are fundamentally quantum mechanical. However, it is possible to employ the use of classical mechanics with a force-field approximation: the energy system is expressed only as a function of the nuclear positions, according with the Born-Oppenheimer approximation. Being an atomistic method, every atom that constitutes the system is modeled as a single material particle and an appropriate potential is introduced to predict the energy associated with the given conformation of the molecule.

Given a conformation vector $\Psi(t) = [\mathrm{pos}_{a_1}, \dots, \mathrm{pos}_{a_n}]^T$, $a_i \in \mathcal{A}$ of a SWCNT $\mathcal{S}$, the potential function V can be written as $V(\Psi(t)) = V_c(\Psi(t)) + V_n(\Psi(t))$, where V_c is the partial term accounting for covalent bonds and V_n for noncovalent interactions. Moreover, the covalent term can be expressed as follows $V_c = V_b + V_a + V_d$, where V_b is the bond stretching potential, V_a is the angle bending potential and V_d is the dihedral interactions potential.

The noncovalent term accounts for the non-bonded electrostatic and van der Waals interactions and can be neglected ([**18**]): only those accounting for bond stretching and angle variation are significant for the system potential. Therefore, the resulting potential energy is approximated as

$$V(\Psi(t)) \simeq V_b(\Psi(t)) + V_a(\Psi(t))\,.$$

Under the assumption of small local deformations, it is adequate to employ the simple harmonic approximation for angles and bonds ([**3**]). In particular, the stretch potential of a single bond b can be expressed as $V_b(\Psi(t)) = k_l/2\left(l(t) - a_{(CC)}\right)^2$ and the bending potential for the couple $c = (b_1, b_2)$ is given by $V_c(\Psi(t)) = k_p/2\left(\arccos\rho_c(t) - 2\pi/3\right)^2$. Hence, the potential of the molecule with the conformation $\Psi(t)$ can be expressed as the summation of the term V_b and V_c for every bond and every couple of adjacent bonds at time t as

$$V = \sum_{b\in\mathcal{B}} V_b + \sum_{c\in\mathcal{C}} V_c.$$

The values of the harmonic parameters are $k_l = 652\,N/m$ and $k_p = 8.76 \cdot 10^{-19}\,Nm$ rad^{-2} ([**15**]).

For every atom a there is an external force that can be described as $f_a(t)$. Thus, the external loads acting on the entire molecule are accounted by the sum

$$L(\Psi(t)) = \sum_{a \in \mathcal{A}} \left(\mathrm{pos}_{a[\Psi]}(t)^T f_a(t) \right) .$$

Finally, the formula for the potential energy and the loading term is:

$$I(\Psi(t)) = V(\Psi(t)) + L(\Psi(t)) .$$

From this formulation it is possible to directly compute the gradient of the scalar potential; for further details refer to [**14**].

2.1.3. *The atomic-scale finite element method.* A state of ground energy corresponds to the equilibrium configuration of a solid. In standard FEM, a continuous solid is partitioned into a finite number of elements, each one with its set of discrete nodes. The energy minimization of the solid is obtained by the appropriate determination of the molecular conformation. Likewise, in molecular mechanics the calculation of the atom positions is based on a similar energy minimization.

Let a molecular system have N atoms (N_C in our case), then the energy stored in the atomic bonds is denoted by $V_{\mathrm{tot}}(\Psi)$ where $\Psi = [\mathrm{pos}_1, \ldots, \mathrm{pos}_n]^T$ is a representation of the conformation vector. The total energy of the system is

$$E_{\mathrm{tot}} = V_{\mathrm{tot}}(\Psi) - \sum_{i=1}^{N} \bar{F} \cdot \Psi_i \,, \tag{1.1}$$

with the external force $\bar{F}$ applied to the $i-$th atom and the state of minimal energy is given by

$$\frac{\partial E_{\mathrm{tot}}}{\partial \Psi} = 0 \,.$$

Let $\Psi^{(0)}$ be an initial guess of equilibrium state and $u = \Psi - \Psi^{(0)}$ the displacement, then the Taylor expansion of E_{tot} around $\Psi^{(0)}$ is

$$E_{\mathrm{tot}} \approx E_{\mathrm{tot}}(\Psi^{(0)}) + \left.\frac{\partial E_{\mathrm{tot}}}{\partial \Psi}\right|_{\Psi^{(0)}} u + \frac{1}{2} u^T \left.\frac{\partial^2 E_{\mathrm{tot}}}{\partial \Psi^2}\right|_{\Psi^{(0)}} u.$$

The combination of the above two formulae yields the equation

$$K(\Psi) u = P(\Psi) \,,$$

where K is the stiffness matrix and P is the non-equilibrium force vector, defined as

$$K = \left.\frac{\partial^2 E_{\mathrm{tot}}}{\partial \Psi^2}\right|_{\Psi^{(0)}} = \left.\frac{\partial^2 V_{\mathrm{tot}}}{\partial \Psi^2}\right|_{\Psi^{(0)}}, \quad P = - \left.\frac{\partial E_{\mathrm{tot}}}{\partial \Psi}\right|_{\Psi^{(0)}} = \bar{F} - \left.\frac{\partial V_{\mathrm{tot}}}{\partial \Psi}\right|_{\Psi^{(0)}}$$

where $\bar{F}$ is the force vector in (1.1).

When no bifurcations are present, the resulting nonlinear system can be solved with iterative methods and it is solved iteratively until P reaches zero. For atomistic interactions with pair potentials, K and P can be obtained from the continuous model as a representation of nonlinear spring elements.

In the application of AFEM to carbon nanotubes, each carbon atom interacts with both first and second nearest neighbour atoms, since those are the most relevant interactions that must be considered ([**13**]). These iterations are the result of the bond angle dependence in the interatomic potential. As a result, ten carbon atoms are considered in each element.

Since the potential has two components, namely the bond and the angle parts, also the stiffness matrix can be decomposed as follows

$$K = K_b + K_a$$

where K_b accounts for V_b and K_a for V_a. The simple analytical form of the potential permits the easy computation of the components of K.

For a given element centered at the i-th atom, only the relation with other nine atom positions is significant. The position of every atom in $\mathbb{R}^3$ is identified by exactly three parameters and, hence, only thirty non-zero elements will appear in every row and every column of the stiffness matrix, since the topological distribution follows the honeycomb pattern. The same applies to the element stiffness matrix and the element non-equilibrium force vector, given by

$$K^e = \begin{bmatrix} \left[\frac{\partial^2 V_{\text{tot}}}{\partial\Psi_1\partial\Psi_1}\right]_{3\times 3} & \left[\frac{\partial^2 V_{\text{tot}}}{\partial\Psi_1\partial\Psi_i}\right]_{3\times 27} \\ \left[\frac{\partial^2 V_{\text{tot}}}{\partial\Psi_i\partial\Psi_1}\right]_{27\times 3} & [0]_{27\times 27} \end{bmatrix}, \quad P^e = \begin{bmatrix} \left[\bar{F} - \frac{\partial V_{\text{tot}}}{\partial\Psi_1}\right]_{3\times 1} \\ [0]_{27\times 1} \end{bmatrix}.$$

The nonlinear resolution of the system $K(\Psi)u = P(\Psi)$ was obtained by the iteration of the four steps that follow:

(i)**:** explicitly compute $P(\Psi^{(i)})$;
(ii)**:** explicitly compute $K(\Psi^{(i)})$;
(iii)**:** solve $u^{(i+1)}$ from $Ku = P$;
(iv)**:** update $\Psi^{(i+1)}$, where $\Psi^{(i+1)} = u^{(i+1)} - \Psi^{(i)}$,

with $i = 0, 1, \ldots, \eta$, until $||P(\Psi^{(\eta)})|| < \varepsilon$, where $\varepsilon = 10^{-6}$ is a predefined tolerance ([**43**]).

2.1.4. *Mechanical properties.* The obtained model allows the analysis of the dependence of length, diameter and elasticity of CNTs on curvature and helicity, compared and validated with the available experimental and computational data ([**14**]).

The results of the simulation of the length at rest of several SWCNTs are plotted as a function of the diameter in Figure 5 for all armchair tubules $(n, 0)$ with $n = 3, \ldots, 11$ and for all zigzag SWCNTs (n, n) with $n = 5, \ldots, 13$.

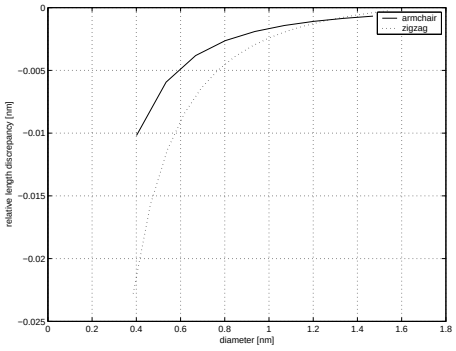

FIGURE 5. Length discrepancy for armchair and zigzag CNTs

In these plots, the length discrepancy is defined as $\delta l = l_g - l_0$, where l_0 is the length on a graphene sheet approximation and l_g is the length at the energy ground for the tube. Length discrepancy is reduced with the increase of the diameter and follows the approximation $\delta l(d) \propto 1/d^2$ for all the tubules, almost vanishing for diameters smaller than 1.2 nm.

Another effect of the curvature is the discrepancy $\delta r = d_h - d_g$ between the estimation of the diameter with the norm of the chirality vector $d_h = 1/\pi||C_h||$ and the real diameter at the energy ground d_g. In Figure 6 the radial discrepancy is plotted for armchair and zigzag nanotubes, showing the smaller is the radius, the larger d_g becomes with respect to d_h.

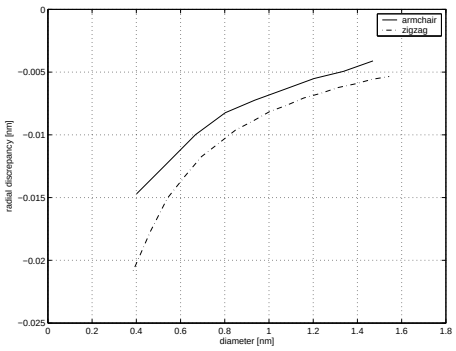

FIGURE 6. Radial discrepancy for armchair and zigzag CNTs

In brief, zigzag nanotubes present a double intensity of the length discrepancy when compared with armchair tubes. Similarly, the former class presents a radial discrepancy whose intensity is approximatively 4/3 of the value obtained for the latter.

This results are comparable with existing investigations ([**66**]).

The first step to describe a structural material is to present its Young's modulus $\bar{E}$ ([**17**]). For a thin rod of isotropic and homogeneous material of cross-sectional area A_0 and of length l_g, $\bar{E}$, is given by $\bar{E} = (l_g F)/(A_0 \Delta l)$, where Δl is the elongation after the load and F is the force applied to the cross-sectional area.

These results are plotted as a function of the diameter for all the armchair tubules $(n, 0)$ with $n = 3, \ldots, 17$ and for all zigzag SWCNTs (n, n) with $n = 5, \ldots, 19$ and $A_0 = \pi(d_g/2)^2$. In the simulations, models with comparable length have been used, with an applied force on the top surface of -0.05 nN per atom in the axial direction of the tubule.

The results for the Young's modulus, obtained under the thin shell assumption, are presented in Figure 7. From the plotted data, the Young's modulus increases with the diameter and tends to the same value of the Young's modulus for graphene sheet as the radius tends to infinity.

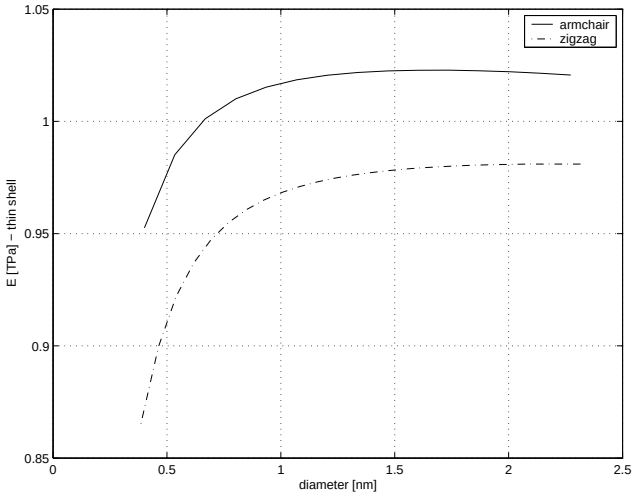

FIGURE 7. Young's moduli for armchair and zigzag nanotubes

In general, the experimentally and computationally predicted Young's moduli for SWCNTs range from 0.5 to 5.5 TPa ([**56**]). The large difference is not only due to the gap between computational and experimental approaches. From the computational point of view, several factors that may strongly influence the results can be summarized as follows: different methods, different molecular parameters, different reference models, different computational conditions. From the experimental point of view, the factors that may diffuse the results can be summarized as: different synthesis of the sample, different measurement techniques, presence of defects and their type, challenging identification of the sample.

Regardless of the reference model, simulations suggest that armchair nanotubes are slightly stiffer than zigzag tubes, but the relative difference is upper bounded by 5% and, thus, can be considered negligible, as noted by several authors ([**5**]) with a wide range of methods. In brief, Young's modulus approximatively does not dependent on chirality.

Poisson's ratio ν is a measure of the tendency of a material, when stretched in one direction, to get thinner in the remaining directions. More precisely, it is the ratio of the

relative contraction strain normal to the applied load and the relative extension strain. In the case of CNTs one has $\nu = -(d_g \Delta l)/(l_g \Delta d)$ where Δl is the elongation after the load and Δd is the difference in diameter after the load. In the numerical simulations, an axial force of -0.05 nN per atom on one surface is used and the minimization does not suffer bifurcations. Figure 8 plots the Poisson's ratio for armchair and zigzag nanotubes.

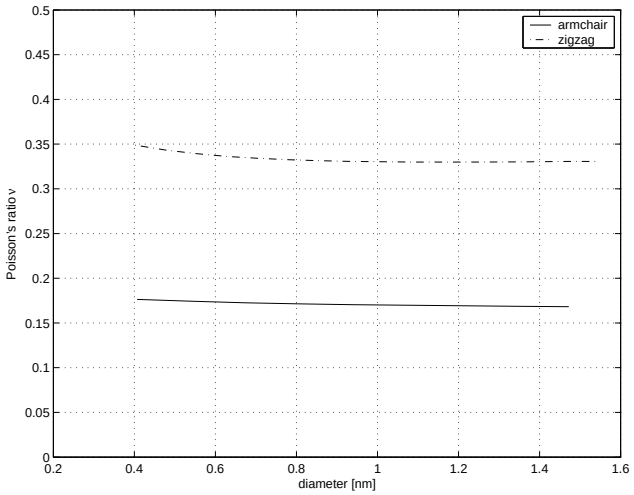

FIGURE 8. Poisson ratio for armchair and zigzag nanotubes

The obtained Poisson's ratios are approximatively not dependent on the radius of the tubules. In fact, narrow tubes have slightly larger Poisson ratios that, however, are upper bounded by 5%. Vice versa, the Poisson's ratio seems to depend in a significant way on the chirality. Indeed, for armchair tubes it's value is $\nu_a = 0.18$, which is almost half of the value $\nu_z = 0.33$ for zigzag tubes. Consequently, it is always necessary to specify the chirality of the tube when the data is analyzed. The values for the Poisson's ratio range from 0.15 ([**36**]) to 0.34 ([**63**]).

The shear modulus G, also called the modulus of rigidity, describes the tendency of a material to shear, that is to deform its shape at constant volume when acted upon by opposing forces. It is defined as shear stress over shear strain. For an isotropic elastic material the Young's modulus $\bar{E}$, the Poisson's ratio ν and the shear modulus G are related as follows $G = \bar{E}/(2 + 2\nu)$. Hence, under the assumption of an isotropic elastic material, Figure 9 plots the results for the obtained shear moduli for armchair and zigzag nanotubes.

Due to the explained difficulties with experimental techniques, there are still a small number of reports on the measured values of shear modulus of SWCNTs. They slightly depend both on chirality and on diameter, as observed in Ref. [**37**]. Theoretical predictions ([**45**]) obtained the average result of 0.5 TPa, which is comparable to the values presented in this work. The lattice-dynamics model permitted the derivation of an analytical expression for the shear modulus, indicating that the shear modulus is about equal to that of graphene for large radii and smaller for narrow tubes ([**37**]). Moreover, it made possible the observation of a weak dependence of the shear modulus on the chirality of the tubes for small radii.

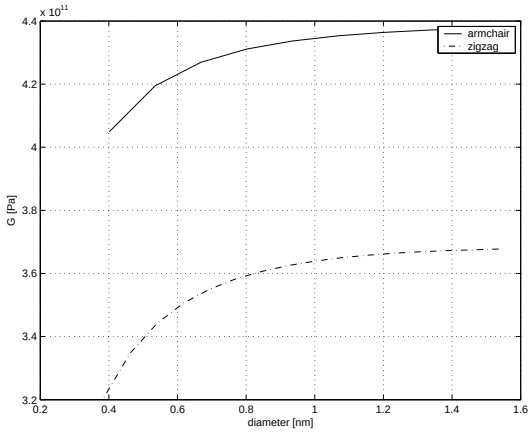

FIGURE 9. Shear moduli for armchair and zigzag nanotubes.

The application of the atomic-scale finite element method with the use of the harmonic approximation for bond stretch and angle bending potentials is presented for the simulation of the mechanical behaviour of carbon nanotubes.

Several simulations have been performed for the numerical results for the Young's moduli, the Poisson's ratios, the shear moduli and the ground energy configurations, that include the evaluation of the lengths and of the diameters at rest applied to a broad range of carbon nanostructures. The obtained parameters and data show agreement with complementary results that come from the experimental data obtained by other authors.

CHAPTER 2

Electric Structure of Single Wall Carbon Nanotubes

We now present the electronic structure of CNTs and the assumptions about their electronic behavior that will lead us to the physical model and to the kinetic equations we concentrate on. We will be dealing only with *metallic* SWCNTs.

Carbon nanotubes are characterized by two types of bond, in analogy with graphene, which exhibits so-called planar sp^2 hybridization. The (s, p_x, p_y) orbitals combine to form in-plane σ (bonding) and σ^* (anti-bonding).

The σ bonds are strong covalent bonds responsible for most of the binding energy and elastic properties of the graphene sheet. The remaining p orbitals interact and create π (bonding) and π^* (anti-bonding) orbitals. The π bonds are perpendicular to the surface of the nanotube and are responsible for the weak interaction between SWNTs in a bundle, similar to the weak interaction between carbon layers in pure graphite ([**19**]). The energy levels associated with the in-plane σ bonds are known to be far away from the Fermi energy in graphite and thus do not play a key role in its electronic properties. In contrast, the bonding π and antibonding π^* bands cross the Fermi level at high-symmetry points in the Brillouin zone of graphene. For a good understanding of the electronic properties of SWNTs, the electronic structure of graphene will be briefly discussed in the next section.

1. From graphene to SWCNT

Graphene is a special semimetal whose Fermi surface is reduced to the distinct points, particular points usually written as K and K', in the hexagonal Brillouin zone. Close to the Fermi energy, the π and π^* bands are nearly linear, in contrast with the quadratic energy-momentum relation obeyed by electrons at band edges in conventional semiconductors. This linear energy-momentum relation of electrons will explain the extremely good conductivity in graphene and bears much importance in the Luttinger-liquid (LL) behavior for low-energy excitations in nanotubes.

The bonding and anti-bonding σ bands are well separated in energy ($> 10eV$). These bands are frequently neglected in semi-empirical calculations as they are too far away from the Fermi level to play a role in the electronic properties of graphene. The remaining two π bands can be simply described with a rather simple tight-binding Hamiltonian, leading to analytical solutions for their energy dispersion and the related eigenstates ([**19**]).

This simple approach can be further extended to study the properties of nanotubes by combining these analytic results with the requirement that the wave functions in tubes must satisfy the proper boundary conditions around the tube circumference. Due to periodic boundary conditions along the circumferential direction of the tube, the allowed wave vectors "around" the nanotube circumference are quantized: they can take only a set of discrete

values. In contrast, the wave vectors along the nanotube axis remain continuous (for infinite tubes). Plotting these allowed vectors for a given nanotube onto the Brillouin zone of graphene generates a series of parallel lines. The length, number, and orientation of these cutting lines depend on the chiral indices (n, m) of the nanotube.

The basic idea behind the *zone-folding approximation* is that the electronic band structure (which is the ε vs. k relation) of a specific nanotube is given by the superposition of the graphene electronic energy bands along the corresponding allowed k lines (k denotes the components along the tube axis of the momentum vectors).

The application of periodic boundary conditions around the tube circumference leads to some restrictions on the allowed wave function quantum phase; depending on the tube symmetry, that is on the chiral vector $C_h = (n, m)$, only two situations can occur ([**19**]): carbon nanotubes are

- metals, when $n - m \equiv 0 \ (mod)\ 3$;
- semiconductors when $n - m \not\equiv 0 \ (mod)\ 3$.

In summary, carbon nanotubes can be metals or semiconductors with an energy gap that depends on the tube diameter and helicity, i.e., on the indices (n, m). This approach is made relatively simple in nanotubes because of the special shape of the graphene Fermi surface and the restriction of the electronic bands to the $(\pi - \pi^*)$ manifold.

Curvature effects have been neglected in this description; carbon nanotubes are not just stripes of graphene but small cylinders. Carbon atoms are placed on a cylindrical wall, a topology that induces several effects different from those of a planar graphene sheet and all these "perturbations" should be considered. For example, for small tubes, the curvature is so strong that some rehybridization among the σ and π states appears. In such a case, the zone-folding picture may fail completely and other calculations should be performed to predict the electronic properties of small diameter nanotubes. For nanotubes with diameters greater than 1 nm, these rehybridization effects are unimportant. Further, symmetry considerations suggest that armchair tubes are less affected by such rehybridization.

When accounting for curvature effects, the only zero-band-gap tubes are the (n, n) armchair nanotubes. Armchair tubes are sometimes labeled type-I metallic tubes, while the others are of type-II. The (n, m) tubes with $n - m = 3l$, where l is a nonzero integer, are tiny-gap semiconductors (the gap being a few tenths of an eV). For the tiny-gap semiconducting nanotubes, the so called secondary gap (due to the curvature) depends on the diameter and the chiral angle, and scales as $1/d_h^2$. This secondary gap is so small that, for most practical purposes, all $n - m \not\equiv 0$ tubes can be considered as metallic at room temperature. Ultrasmall radius single-wall carbon nanotubes (diameter of about 4 $\mathring{A}$) have many unusual properties such as superconductivity ([**62**]). The properties of these ultrasmall tubes have already been extensively investigated.

2. Transport properties

In the absence of scattering, i.e. when transport is *ballistic*, the resistance of a metallic SWCNT is given by Landauer's equation. This resistance is a contact resistance arising from the mismatch of the number of conduction channels in the CNT (= 2) and the macroscopic metal leads (= 10^6 for a 1 μm metal lead). In addition to this quantum mechanical contact resistance, there are other sources of contact resistance, such as that

produced by the existence of interface barriers, or poor coupling between the CNT and the leads. These types of resistance are very important and can dominate electrical transport in nanotubes. Other quantum effects which have to be considered, in order to obtain a reliable physical model, are electron-electron scattering and electron-phonon scattering processes. It is evident, both from the experimental and the theorical point of view, that these processes can not be neglected, otherwise the resulting predictions would always be wrong.

We will discuss such processes in a more quantitative way in the next section and in Chapter 5.

3. Quantitative model

Electrical properties of SWCNTs arise from the confinement of electrons, which allows motion in only two directions, and the requirements for energy and momentum conservation. These three constraints lead to a reduced phase space for scattering processes ([**9**]).

As long as we restrict our interest to low energies (i.e., a few hundred meV from the Fermi surface E_F) the band-structure of a metallic nanotube can be approximated by two sets of bands with a *linear* dispersion intersecting at k_F and $-k_F$ (Figure 1).

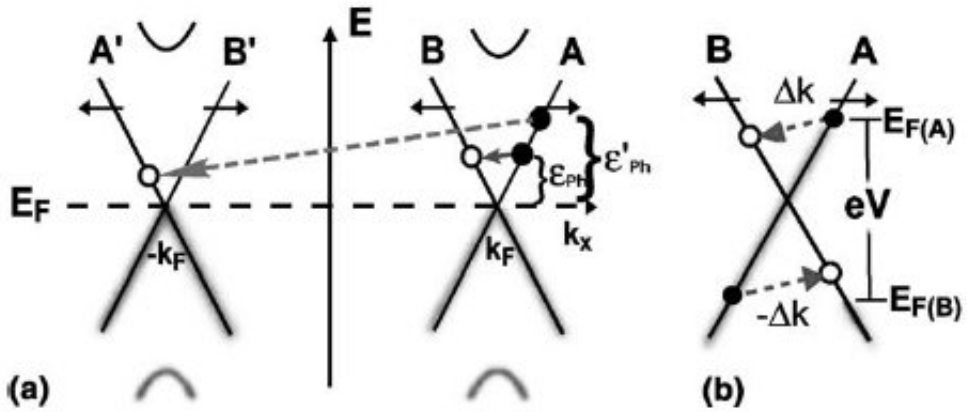

FIGURE 1. Band-structure of a metallic nanotube at low energies (a) and under an applied bias voltage (b) ([**9**])

Electrons with positive k move towards the right, while electrons with negative k move to the left. The situation is similar for semiconducting CNTs, except for the fact that the two bands do not cross at E_F, but a band-gap $\varepsilon_{Gap} = (4\hbar v_F/3d_h)$, is present; we will not discuss this situation here.

Electron energies $\varepsilon(k) - E_F$ near the Fermi level E_F are well approximated by the linear dispersion relations:

$$\varepsilon_i(k) = \hbar v_i k, \qquad i = 1, 2, \tag{2.1}$$

with positive $v_1 = +v_F$ and negative $v_2 = -v_F$ Fermi velocities; $\hbar$ is the reduced Plank constant. Electrons in these states can be considered as equivalent carriers in the transport model.

It is thus sufficient to introduce two phase-space distribution functions $f_i = f_i(t, x, \varepsilon)$ with i referring to right ($i = 1$) and left moving ($i = 2$) electrons. Distribution functions $f_i(t, x, \varepsilon)$ depend on time t, the position variable x and the energy $\varepsilon = \varepsilon_i(k)$.

All further semiconductorlike energy sub-bands of the nanotube are neglected, which is valid for electrons with energies $|\varepsilon| < 2\hbar v_F/d_h$. Assuming a tube diameter $d_h = 2\,nm$, this energy bound is approximately $0.55\,eV$.

The temporal evolution of the distribution functions $f_i = f_i(t, x, \varepsilon)$ is governed by the following Boltzmann equations

$$\frac{\partial f_i}{\partial t} + v_i \frac{\partial f_i}{\partial x} - e_0 E v_F \frac{\partial f_i}{\partial \varepsilon} = \mathcal{C}_i, \qquad i = 1, 2, \tag{2.2}$$

where e_0 denotes the electron charge (considering the negative value) and E the electric field along the tube axis. Collision operators $\mathcal{C}_i$ on the right hand side of equation (2.2) determine the temporal changes of f_i due to electron scattering.

Equation (2.2) represents a kinetic semiclassical *Boltzmann equation* and it is the starting point of many simulations regarding CNTs electronics. Usually, an important difference between models is the explicit formulation of the collision operators $\mathcal{C}_1$ and $\mathcal{C}_2$. In almost all formulae which can be found in the literature, collision operators are always given by the sum of two terms: one taking into account electron-electron scattering processes and the other accounting for electron-phonon scattering.

Electron-electron interaction is taken into account by scattering of electrons with acoustic phonons in terms of elastic back-scattering processes. The corresponding collision operator usually takes the form:

$$\mathcal{C}_i^{ac} = \frac{v_F}{l_{ac}}(f_j - f_i) \tag{2.3}$$

with $j \neq i$ and the acoustic mean free path l_{ac}.

It should be mentioned that the effect of electron scattering at impurities can be included by substituting l_{ac} with the elastic MFP l_e:

$$\frac{1}{l_e} = \frac{1}{l_{ac}} + \frac{1}{l_{im}},$$

where l_{im} denotes the MFP due to impurity scattering.

In one of the first models about CNTs electronics ([**65**]), authors considered electron-phonon scattering contribution as given by two terms; the first representing back scattering from phonons:

$$\mathcal{C}_i^{pb} = \frac{v_F}{l_{pb}}\big[(1 - f_i)f_j^+ - f_i(1 - f_j^-)\big]$$

and the other representing forward scattering from phonons:

$$\mathcal{C}_i^{pf} = \frac{v_F}{l_{pf}}\big[(1 - f_i)f_i^+ - f_i(1 - f_i^-)\big].$$

Here, $f_i^{\pm}$ are f_i evaluated at $\varepsilon \pm \hbar\omega$, i.e. $f_i^{\pm} = f_i(t, x, \varepsilon \pm \hbar\omega)$, where $\hbar\omega$ is the phonon energy quantum; l_{pb} is the distance an electron travels before back scattering, once the phonon emission threshold has been reached and, similarly, l_{pf} is the distance an electron travels before forward scattering.

In the presented model, authors have assumed that the heat generated in the tube escapes sufficiently quickly to avoid raising the lattice temperature too high; this relies on the assumption of phonons in equilibrium at a fixed lattice temperature. A simple estimate of nanotube's thermal conductivity indicates that it is unlikely that all of the heat could be transmitted through the contacts. Both acoustic phonons and substrate are thermalized at room temperature, and are acting as a thermal bath.

Such model is not valid if the tube is suspended between the two electrodes. In this case, indeed, the acoustic phonons are not thermalized with the environment and their occupation should be determined self-consistently. In [**49**], simple macroscopic approaches are presented to include the self-heating effect in suspended SWCNT's. Nonequilibrium effects of the optical phonon system are incorporated by considering an effective temperature for optical phonons different from the lattice temperature. However, studies of strongly coupled electron-phonon systems prove that electrons' transport properties are strongly influenced by the nonequilibrium shape of the phonon distributions. If strong kinetic effects occur, transport models based on equilibrium phonon distributions with effective temperatures become very inaccurate.

If high electric fields are applied, the current is essentially limited by inelastic interactions of electrons with optical phonons. In [**6**] and [**39**] electrons transport in SWCNTs under high bias is entirely treated at the kinetic level and include the nonequilibrium dynamics of optical phonons.

A transport model for metallic SWCNT's is proposed, based on a coupled system of semiclassical Boltzmann equations for both electrons and phonons. In this way, the influence of the nonequilibrium behavior of the phonon distributions on the electron transport is taken into account dynamically.

Let $g_\eta = g_\eta(t, x, q)$ be phonons occupations, where q is the 1D wave vector and η range over the different phonon modes taken into account. Phonon distributions evolve according to

$$\frac{\partial g_\eta}{\partial t} + \nu_\eta \frac{\partial g_\eta}{\partial x} = \mathcal{D}_\eta, \tag{2.4}$$

with phonon velocities $\nu_\eta = \partial_q \omega_\eta$ and phonon collision operators $\mathcal{D}_\eta$. In this model, phonon dispersion relations are treated according to the Einstein approximation, by assuming constant phonon energies $\hbar\omega_\eta$; however, nonzero phonon velocities ν_η are considered in eq. (2.4) to incorporate the effect of spatial diffusion of optical phonons ([**6**]).

Electron-phonon interactions, i.e. right hand side of equations (2.2) and (2.4), are obtained by using Fermi's golden rule. The application of Fermi's golden rule leads to:

$$\begin{aligned} s_{ij} &= \frac{2\pi}{\hbar} \sum_\eta \int |c^\eta_{jk+q,ik}|^2 \times \\ &\times \left\{ g_\eta(q)\delta[\varepsilon_i(k) - \varepsilon_j(k+q) + \hbar\omega_\eta(q)] + \right. \\ &+ \left. [g_\eta(-q) + 1]\delta[\varepsilon_i(k) - \varepsilon_j(k+q) - \hbar\omega_\eta(-q)] \right\} dq \end{aligned}$$

for the scattering rate of electrons in the state (i, k) due to absorption and emission of phonons with wave vectors q and $-q$; electron-phonon coupling (EPC) is determined by the quantity $c^\eta_{jk',ik}$ (refer to [**8**] for a discussion on the validity and of these assumptions to SWCNTs).

In [**39**], two optical phonon modes $\eta = 1, 2$, are considered; they represent K (or zone-boundary) phonons and longitudinal optical Γ phonons, respectively. In [**6**], one more scattering process is taken into account, thus (the distribution evolutions of) three phonon modes $\eta = 1, 2, 3$ are considered: they refer to K phonons, longitudinal optical Γ phonons and transverse optical Γ phonons, respectively. Explicit formulae for collision operators of both electrons and phonons will be given in the last chapter.

Calculations show that the optical phonons are driven far from equilibrium at high applied bias, especially at the contact boundaries. This nonequilibrium behavior of the optical phonons is found to influence significantly the electron transport.

From a mathematical point of view, equations (2.2) and (2.4) together constitute a system of Conservation Laws (CL) with source terms, commonly known as Balance Laws (BL). Distribution functions $f_i(t, x, \varepsilon)$ and $g_\eta(t, x, q)$ depend both on two phase-space variables. Therefore, the kinetic transport model defined by Boltzmann equations (2.2) and (2.4) can be solved very efficiently with the help of a deterministic solver, thus avoiding statistical methods (Monte Carlo methods are usually adopted for this kind of computations).

We will present the problem numerical approximation and simulations results in the last chapter but first we will review, in the next two chapters, the general mathematical theory regarding Balance Laws and then their numerical treatment.

CHAPTER 3

Balance Laws

The general setting of our investigation is that of hyperbolic *Conservation Laws* (CLs) with source terms, commonly known also as *Balance Laws* (BLs). Conservation Laws are time dependent systems of Partial Differential Equations having, in general, the following form:

$$\frac{\partial}{\partial t}u(t,x) + \frac{\partial}{\partial x}f(x,u(t,x)) = 0. \tag{3.1}$$

Here $u : \mathbb{R}\times\mathbb{R}^d \to \mathbb{R}^m$ is a vector of conserved quantities and $f : \mathbb{R}^m \to \mathbb{R}^m$ is usually known as the flux function. In BLs, the time evolution of the conserved quantities is balanced by a source (or eventually sink) term on the right-hand side of the equation:

$$u_t + f_x(x,u) = s(x,u). \tag{3.2}$$

Without loss of generality, we can restrict to the study of the homogeneous case, i.e. when f and s depend only on u and not on x (see [**22**]).

In the monodimensional case, when $d = 1$, a system of CLs is said *hyperbolic* if the $m\times m$ Jacobian matrix A of f

$$A(u) = \frac{\partial}{\partial u}f(u)$$

has m real eigenvalues and a complete set of eigenvectors (A is diagonalizable with all real eigenvalues); the system is strictly hyperbolic if the eigenvalues are all distinct. The general definition for the multi-dimensional case will be given later.

Equations of type (3.1) are called conservation laws since they represent the conservation of the *extensive* quantity $\int_{x_1}^{x_2} u\,dx$ when the incoming flux and outgoing flux balance each other; indeed

$$\frac{\partial}{\partial_t}\int_{x_1}^{x_2} u\,dx = -\int_{x_1}^{x_2} f_x(u)\,dx = -(f(u(x_2) - f(u(x_1))).$$

We present a general introduction on CLs which will help to understand the more complicated behaviour of balance laws. The resume presented in this chapter and in the next one is inspired by the books ([**42, 22, 28**]).

1. The scalar equation

Many of the characterizing properties of general CLs can already be found in the scalar mono dimensional case, i.e. $d = m = 1$. For this reason, and for simplicity, we start by first describing this simpler, but meaningful, case.

1.1. The advection equation. We start from the simplest of all CLs, the *linear advection* equation:

$$u_t + \lambda u_x = 0, \tag{3.3}$$

where λ is a real parameter; we consider such equation to be defined on the domain $-\infty < x < +\infty$ and $t \geq 0$ and we assume initial condition:

$$u(0, x) = u_0(x).$$

It is easy to see that the solution to this problem is a *self-similar* equation:

$$u(t, x) = u(x - \lambda t),$$

which is the initial data moving unchanged in shape to the left ($\lambda < 0$) or to the right ($\lambda > 0$) with velocity λ.

From this very simple case we can already infer one of the peculiar properties of hyperbolic PDEs: they admit discontinuous "solutions": if the initial value u_0 is a discontinuous function, so will be the general solution (as we said, it is the same function, moving unchanged). Moreover, it is easy to see that the solution u is constant along particular curves, called *characteristics* of the equation; for the advection equation, characteristics are the rays $x - \lambda t = x_0$, which can be found solving

$$\begin{cases} x'(t) = \lambda, \\ x(0) = x_0. \end{cases}$$

Along such curves, the value of u is constant; indeed:

$$\frac{\partial u}{\partial t}(t, x(t)) = u_t + u_x \, \lambda = 0.$$

From this simple case we can also see that the solution u at a point $(\bar{t}, \bar{x})$ depends only on the value of the initial data u_0 at the point $\bar{x}_0 = \bar{x} - \lambda \bar{t}$, which lies on the same characteristic line of $(\bar{t}, \bar{x})$ at $(0, \bar{x}_0)$. We could change the initial data at any points other than $\bar{x}_0$ without affecting the solution $u(\bar{t}, \bar{x})$. The set $\mathcal{D}(\bar{t}, \bar{x}) = \{\bar{x}_0\}$ is called the *domain of dependence* of the point $(\bar{t}, \bar{x})$. Here this domain consists of a single point; for a system of equations this domain is typically an interval and a fundamental fact about hyperbolic equations is that it is always a bounded interval. The size of this set usually increases with time, but we have a bound of the form $\mathcal{D} \subset \{x : |x - \bar{x}| < \lambda_{max} \, \bar{t}\}$ for some value λ_{max}.

Conversely, initial data at any given point x_0 can influence the solution only within some cone $\{x : |x - x_0| < \lambda_{max} \, t)\}$ of the $x - t$ plane. This region is called the *range of influence* of the point x_0. We summarize this by saying that hyperbolic equations have *finite propagation speed*; information can travel with speed at most λ_{max}. This has important consequences in developing numerical methods.

1.2. The non linear case. In the more general non linear case, we can assume f is differentiable (this is true in all relevant physical models) so we can rewrite (3.1) in the quasi-linear form:

$$u_t + \lambda(u) u_x = 0, \tag{3.4}$$

where $\lambda(u) = f'(u)$. Characteristics can be found, as in the previous case, solving

$$\begin{cases} x'(t) = \lambda(u), \\ x(0) = x_0. \end{cases}$$

The solution u is again constant along such curves:

$$\frac{\partial u}{\partial t}(t, x(t)) = u_t + u_x \, \lambda(u) = 0.$$

We can see that also in this case the solution depends only on the value of u_0 on a single point x_0; in particular, it is constant along the characteristic. This is a very useful fact because it can be used to compute the general solution, as in the linear case. The value of u at (t, x) is equal to $u_0(x_0)$, where x_0 is the solution of:

$$x = x_0 + t\lambda(u_0(x_0)).$$

From this simple example we understand another very important fact. We see why a hyperbolic PDE admits discontinuous "solutions": at the point (t, x) the solutions depends only on the one value $u_0(x_0)$, so it is clear that global spatial smoothness is not required for this construction of $u(t, x)$. We can, for this reason, define a global solution even if the initial data is not a smooth function.

We can also see that discontinuities travel only along characteristics, which means they do not affect the solution in other points.

If u_0 is not differentiable at some point then $u(t, x)$ is no longer a classical solution of the differential equation everywhere. However, this function does satisfy the integral form of the conservation law and the integral form does make sense even for nonsmooth u. Recall that the integral form is more physical than the differential equation, which was derived from the integral form under the additional assumption of smoothness. It thus makes perfect sense to accept this generalized solution to solve conservation laws.

Unlike in the linear case, in the non linear case it is also possible that, even if initial data is smooth, a discontinuity may create during the time evolution of the solution. From any point x_0, the solution moves along the characteristic with speed equal to $\lambda(u_0(x_0))$, which remain constant along the characteristic but varies depending on x_0. So if there exist two points $x_1 < x_2$ having angular coefficient of the characteristic line $m_1 = 1/\lambda(u_0(x_1)) < m_2 = 1/\lambda(u_0(x_2))$, then the two lines will intersect at a finite time T_b (called *break* time). In general, we see that two characteristics intersect at time $\bar{t}$ if

$$\bar{t}(\lambda(u_0(x_1)) - \lambda(u_0(x_1)) = x_2 - x_1.$$

Thus, unless the function $\lambda(u_0(\cdot))$ is monotonically non-decreasing, in which case this equation has no positive solution $\bar{t}$, we cannot define a classical solution u for all time $t > 0$. It is possible to determine the critical time T_b up to which a classical solution exists and can be constructed by the method of characteristics. For convex flux functions, time T_b can be computed by:

$$T_b = -\frac{1}{\min_{x \in \mathbb{R}} \left(\lambda(u_0'(x))\right)}.$$

Beyond this point there is no classical solution of the PDE, and the solution we wish to determine becomes discontinuous. The classical solution ceases to exist but a weak solution can still be defined. At the breaking point, a discontinuity, or *shock* forms.

If we only considered classical solutions, we would not be able to solve our problem in the correct way or at least in complete way. As we already said, the differential form is not the truly physically relevant one, the integral form being the correct one. It is known that the latter admits more solutions than only the differentiable ones. The *variational* or *weak* formulation of the equation has to be considered in order to recover the full behavior of the problem. After we state the problem in the weak form, we can consider more general solutions, which means we require less regularity for the solution u. Following the standard procedure, let $\phi \in \mathcal{C}_0^1(\mathbb{R}^+ \times \mathbb{R})$ be a test function, where $\mathcal{C}_0^1$ is the space of continuously differentiable functions with compact support. If we multiply $u_t + f_x = 0$ by ϕ and then integrate over space and time, we obtain

$$\int_0^\infty \int_{-\infty}^\infty \big(u_t + f_x(u)\big)\phi \, dx \, dt$$

which, after a formal integration by parts, yields

$$\int_0^\infty \int_{-\infty}^\infty \big(u\phi_t + f(u)\phi_x\big) \, dx \, dt \;\; = \;\; -\int_{-\infty}^\infty u(0,x)\phi(0,x) \, dx = \tag{3.5}$$

$$-\int_{-\infty}^\infty u_0(x)\phi(0,x) \, dx. \tag{3.6}$$

We remark that (3.5) makes sense if $u, u_0 \in \mathcal{L}^\infty_{loc}(\mathbb{R} \times \mathbb{R}^+)$, where L^∞_{loc} is the space of locally bounded measurable functions.

It is easy to verify that this is a true generalization of the classical notion of solution: if u is a regular function we can recover the differential formulation.

Unfortunately, now we have too many solutions available ([**42, 28**]): weak solutions are not, in general, unique and one needs to find a way to identify the only physically correct one.

1.2.1. *R-H condition.* First of all, we could check whether the solution propagates at the correct velocity or not. The speed of propagation can be determined by conservation; let $[h] = h^- - h^+$ the value of the jump. Easy calculations show that the following relation:

$$\int_\Gamma \big(- [u]x'(t) + [f(u)]\big)\phi \, dt = 0$$

holds for every smooth curve Γ across which u has a jump discontinuity; in this case, $[u] = u^- - u^+$, with $u^- = u^-(t, x(t))$ and $u^+ = u^+(t, x(t))$).

Since ϕ was arbitrary, we obtain

$$[u]\xi = [f(u)] \tag{3.7}$$

at each point on Γ, so everywhere along the discontinuity. Here, $\xi = x'(t)$ is the speed of the discontinuity. Relation (3.7) is called the *Rankine-Hugoniot condition.* It holds, formally unchanged for systems of equations. For a linear system with $f(u) = A(u)$, for example, we obtain

$$A[u] = \xi[u]$$

i.e., $u_l - u_r$ must be an eigenvector of the matrix A and ξ its associated eigenvalue. For a linear system, these eigenvalues are the characteristic speeds of the system. Thus discontinuities can propagate only along characteristics, a fact that we have already seen for the scalar case.

1.2.2. *Retrieving the correct solution.* Ways to select the correct physical solution have been widely studied. A first guess is to consider an approximation of the non smooth initial data u_0 by a sequence of smooth functions $u_0^\varepsilon(x)$ s.t.

$$\|u_0 - u_0^\varepsilon\|_1 < \varepsilon$$

as $\varepsilon \to 0$. We can than take the solution to the problem as $u = lim_{\varepsilon\to 0} u_0^\varepsilon$. This method could be useful for some particular situations but in general it will not. One problem is, e.g., that for nonlinear problems singularity may develop during the time evolution also when smooth initial data are present and this method would fail in that case.

The two most important and useful approaches derive, instead, from physical motivations: the first, called *vanishing viscosity* method, comes from the theory of thermoviscoelasticity, where purely hyperbolic equations (i.e. adiabatic thermoelasticity) may be physically meaningful only as a limiting case of general thermoelasticity with viscosity and heat conductivity tending to zero. It is this general philosophy that underlies the vanishing viscosity approach. The other method is based on *entropy conditions*; they are called in this way since they are conditions, a solution has to satisfy, deriving implicitly or explicitly from the second principle of thermodynamics.

A conservation law (3.3) should in general be considered as an approximation to the advection diffusion equation

$$u_t + \lambda u_x = \varepsilon u_{xx} \tag{3.8}$$

for very small $\varepsilon > 0$. If we now let $u^\varepsilon(t, x)$ denote the solution of this equation with initial data $u_0(x)$, then $u^\varepsilon \in \mathcal{C}^\infty((0, \infty) \times (-\infty, \infty))$ even if $u_0(x)$ is not smooth since (3.8) is a parabolic equation. We can again take the limit of $u^\varepsilon(t, x)$ as $\varepsilon \to 0$ and will obtain a physically relevant solution, often referred to as the vanishing viscosity solution.

This method is hard to work with, and a variety of other conditions, the *entropy conditions*, that can be applied directly in order to check whether or not a weak solution is physically admissible have been developed instead.

We resume here the most important and commonly used of them:

entropy condition I: a discontinuity propagating with speed ξ given by (3.7) satisfies the entropy condition if

$$f'(u_l) > \xi > f'(u_r), \tag{3.9}$$

where u_l and u_r are the values at left and right of the discontinuity. Note that $f'(u)$ is the characteristic speed.

A more general form of this condition, due to Oleinik, applies also to nonconvex scalar flux functions f:

entropy condition II: $u(t, x)$ is the entropy solution if all discontinuities have the property that

$$\frac{f(u) - f(u_l)}{u - u_l} > \xi > \frac{f(u) - f(u_r)}{u - u_r} \tag{3.10}$$

for all u between u_l and u_r. For convex f, this requirement reduces to the previous one.

Another form of the entropy condition is based on the spreading of characteristics in a rarefaction fan. If $u(t, x)$ is an increasing function of x in some region, then characteristics

spread out if $f'' > 0$. The rate of spreading can be quantified, and gives the following condition, also due to Oleinik:

entropy condition III: $u(t,x)$ is the entropy solution if there is a constant $E > 0$ such that for all $a > 0$, $t > 0$ and $x \in \mathbb{R}$

$$\frac{u(t,x+a) - u(t,x)}{a} < \frac{E}{t}. \tag{3.11}$$

The form of (3.11) may seem unnecessarily complicated, but it turns out to be easier to apply in some contexts. In particular, this formulation has advantages in studying numerical methods.

Yet another approach to the entropy condition is to define an *entropy function* $\eta(u)$ for which an additional conservation law holds for smooth solutions that becomes an inequality for discontinuous solutions. Suppose some function $\eta(u)$ satisfies a conservation law of the form

$$\eta_t(u) + \psi_x(u) = 0 \tag{3.12}$$

for some *entropy flux* ψ. Then, for smooth u, we obtain

$$\eta'(u)u_t + \psi'(u)u_x = 0. \tag{3.13}$$

Recall that a general conservation law can be written as $u_t + f'(u)u_x = 0$. Multiply this by $\eta'(u)$ and compare with (3.13) to obtain

$$\psi'(u) = \eta'(u)\, f'(u). \tag{3.14}$$

For a scalar conservation law this equation admits many solutions $(\eta(u), \psi(u))$. For a system of equations η and ψ are still scalar functions, but now (3.14) reads

$$\nabla\psi(u) = f'(u)\nabla\eta(u),$$

which is a system of m equations for the two variables η and ψ. If $m > 2$, this may have no solutions.

An additional condition we place on the entropy function is that it be convex. The entropy $\eta(u)$ is conserved for smooth flows by its definition; for discontinuous solutions, however, the manipulations performed above are not valid. The final form of the entropy condition, called the entropy inequality, reads ([**42**]):

entropy condition IV: the function $u(t,x)$ is the entropy solution if, for all convex entropy functions η and corresponding entropy fluxes ψ, the inequality

$$\eta_t(u) + \psi_x(u) \leq 0. \tag{3.15}$$

is satisfied in the weak sense. This formulation is also useful in analyzing numerical methods. If a discrete form of this entropy inequality is known to hold for some numerical method, then it can be shown that the method converges to the entropy solution.

Just as for the conservation law, an alternative weak form of the entropy condition can be formulated:

$$\int_0^\infty \int_{-\infty}^\infty \phi_t(t,x)\eta(u(t,x)) + \phi_x(t,x)\psi(u(t,x))\, dx\, dt \leq -\int_{-\infty}^\infty \phi(0,x)\eta(u(0,x))\, dx \tag{3.16}$$

for all $\psi \in \mathcal{C}_0^1(\mathbb{R}\times\mathbb{R})$ with $\psi(t,x) > 0$ for all t and x.

1.2.3. *The Riemann Problem.* The so called Riemann Problem is a special case of Cauchy Problem for CLs: it is a CL together with piecewise constant initial data having a single discontinuity:

$$u_0(x) = \begin{cases} u_l & x < 0, \\ u_r & x > 0. \end{cases}$$

For the linear advection equation the behavior of the solution is the standard translation. For nonlinear equations many different behavior are possible, instead. As a significant example, if we take into account the *Burger equation*, which is a conservation law with flux $f(u) = u^2/2$, wee have that the form of the solution depends on the relation between u_l and u_r:

`CASE I:` $u_l > u_r$.

In this case there is a unique weak solution:

$$u(x) = \begin{cases} u_l & x < \xi t, \\ u_r & x > \xi t, \end{cases}$$

where

$$\xi = \frac{u_l + u_r}{2}$$

is the shock speed (the speed at which the discontinuity travels). In this case characteristics go *into* the shock; it is easy to see that this is also the vanishing viscosity solution.

`CASE II:` $u_l < u_r$.

In this case there are infinitely many weak solutions. One of these is the previous function, in which the discontinuity propagates with speed ξ. Note that characteristics now go *out* of the shock and that this solution is not stable to perturbations: if the data is smeared out slightly, or if a small amount of viscosity is added to the equation, the solution changes completely.

Another weak solution is the rarefaction wave

$$u(t,x) = \begin{cases} u_l & x < u_l t \\ x/t & u_l t \le x \le u_r t \\ u_r & x > u_r t. \end{cases} \tag{3.17}$$

This solution is stable to perturbations and is in fact the vanishing viscosity generalized solution. There are infinitely many solutions of this type: indeed,

$$u(t,x) = \begin{cases} u_l & x < s_m t \\ u_m & s_m t \le x \le u_m t \\ x/t & u_m t \le x \le u_r t \\ u_r & x > u_r t \end{cases} \tag{3.18}$$

is a weak solution for any u_m with $u_l < u_m < u_r$ and $\xi_m = (u_l + u_m)/2$.

2. Systems of equations

2.1. Linear systems. We now consider a linear system of equations in one space dimension:

$$\begin{cases} u_t + A u_x = 0, \\ u(0,x) = u_0(x), \end{cases} \tag{3.19}$$

where A is a constant coefficient matrix. Recalling the definition given before, the system is said hyperbolic if A can be diagonalized with real eigenvalues; we have

$$A = R\Lambda R^{-1} \tag{3.20}$$

where $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_m)$ and $R = [r_1|\dots|r_m]$ is the matrix of eigenvectors.

In the linear case it is possible to solve the system in a similar way respect to the scalar case: it is possible to compute the propagation directions of the initial data. This can be done via a change of variables, using the so called *characteristic variables*:

$$v = R^{-1}u. \tag{3.21}$$

Multiplying the first equation in (3.19) by R^{-1} we obtain:

$$R^{-1}u_t + R^{-1}Au_x = 0;$$

using (3.20) and since R and R^{-1} are constant, we have

$$v_t + \Lambda v_x = 0.$$

Since Λ is a diagonal matrix, the original system decouples into m independent scalar advection equations

$$(v_p)_t + \lambda_p (v_p)_x = 0 \qquad p = 1, \dots, m.$$

Each of these equations has well known solutions $v_p(t,x) = v_p^0(x - \lambda_p t)$ where v_0 is the initial data given by

$$v^0(x) = R^{-1}u_0(x).$$

We can now go back to the physical variables u by $Rv = u$ to obtain the solution

$$u(t,x) = \sum_{p=1}^{m} v_p(t,x) r_p = \sum_{p=1}^{m} v_p(0, x - \lambda_p t) r_p. \tag{3.22}$$

It is readily seen that solution $u(t,x)$ depends only on the value of the initial data at the points $x - \lambda_p t$, for $p = 1, \dots, m$, so $\mathcal{D}(t,x) = \{\bar{x} = x - \lambda_p t \mid p = 1, \dots, m\}$. The curves $x = x_0 + \lambda_p t$ are the *characteristics of the p-th family* (or $p-$characteristics). Coefficients $v_p(t,x)$ of r_p in the eigenvectors expansion (3.22) of u are constant along any $p-$characteristic. The solution can be viewed as a superposition of m waves, each of which is advected independently with no change in shape, propagating at speed v_p along the direction r_p.

2.2. Non linear systems. In the nonlinear case we have that $A(u)$ can be diagonalized, at least in some range of u where the solution is defined and differentiable.

We can define characteristics as in the linear case: there are m characteristic curves through each point

$$\begin{cases} x'(t) = \lambda_p(u(t, x(t)), \\ x(0) = x_0. \end{cases} \tag{3.23}$$

Characteristic method is not as useful as in the previous case since λ_p depends on u so we can no more use this method to solve the problem; a more complicated coupled system should be used instead but this is not an effective approach.

Characteristics are, anyway, useful where the solution is smooth; for example, linearizing the system at a constant state $\bar{u}$ one obtain a constant coefficient system with the Jacobian

matrix frozen at $A(\bar{u})$; it is then possible to have a local approximation of the solution and also obtain information on the (local) behavior of "small" discontinuities (see [**42**]).

Recall that for linear systems singularities propagate only along characteristics. For nonlinear problems this is not the case, as we have already seen for nonlinear scalar equations. The Rankine-Hugoniot jump condition (3.7):

$$f(u_l) - f(u_r) = \xi(u_l - u_r)$$

must be satisfied for a propagating discontinuity; here ξ is the propagation speed.

2.3. The Riemann Problem. For a constant coefficient linear system, the Riemann problem can be explicitly solved; the solution is well known and given by (3.22). We will see shortly that the solution to a nonlinear Riemann problem has a simple structure, quite similar to the structure of the linear solution.

For the Riemann problem, notation in (3.22) can be simplified decomposing $u_l = \sum_{p=1}^{m} \alpha_p r_p$ and $u_r = \sum_{p=1}^{m} \beta_p r_p$ in the characteristic base: if we let $P(t,x)$ be the maximum value of p for which $x - \lambda_p t > 0$, then

$$u(t,x) = \sum_{p=1}^{P(t,x)} \beta_p r_p + \sum_{p=P(t,x)+1}^{m} \alpha_p r_p.$$

The jump $u_r - u_l$ cannot propagate as a single discontinuity with any speed without violating the Rankine-Hugoniot condition. We can view "solving the Riemann problem" as finding a way to split up the jump into a sum of jumps each of which can propagate at an appropriate speed with the Rankine-Hugoniot condition satisfied.

For nonlinear systems the Riemann problem can be solved in much the same way: we can attempt to find a way to split this jump up as a sum of jumps, across each of which the property described before does hold, although life is complicated by the fact that we may need to introduce rarefaction waves as well as shocks.

2.4. Multidimensional problems. Recall the definition of hyperbolicity for a system in d dimensions

$$\frac{\partial u}{\partial t} + \sum_{j=1}^{d} \frac{\partial f_j}{\partial x_j}(u) = 0, \tag{3.24}$$

where $x \in \mathbb{R}^d$ and $f_j \in \mathbb{R}^m$. For $\omega \in \mathbb{R}^d$, let

$$A(u,\omega) = \sum_{j=1}^{d} \frac{\partial f_j}{\partial x_j} \cdot \omega_j.$$

A system of conservation laws is called hyperbolic if for any direction ω, with $|\omega| = 1$, matrix $A(u,\omega)$ has m real eigenvalues

$$\lambda_1(u,\omega) \leq \ldots \leq \lambda_m(u,\omega)$$

and a complete set of eigenvectors $r_j(u,\omega)$.

Theorical aspects of the multidimensional case are much more complex than those for the one dimensional case or for the scalar case; still nowadays they are not fully understood. A few results are available for $d = 2, 3$ for particular types of equations while not much is known

for d greater than 3. It is possible to extend the definition of characteristics, which are called characteristic (hyper) planes in this context and it is possible to extend also the method of characteristics to find solutions of linear problems and at least locally in the nonlinear cases. In particular, for nonlinear systems, it is still possible to define traveling "plane waves", transporting the solution unchanged. As a complete reference for the multidimensional case and related issues we suggest [**28**].

2.5. Source terms. The study of generic Balance Laws is harder than that of Conservation Laws. Some theoretical results are available for the scalar equation but very few is known for systems in one or more dimensions, apart from simple special cases. As a complete reference one can refer to [**22**].

Here we will show an example of solution just for the linear constant coefficient case. Consider the 1D scalar advection equation with source term:

$$u_t + \lambda u_x = s(u). \tag{3.25}$$

Let

$$\begin{cases} x'(t) = \lambda, \\ x(0) = x_0 \end{cases}$$

and

$$v(t) = u(t, x(t)).$$

Then:

$$v'(t) = u_t(t, x(t)) + \lambda u_x(t, x(t)) = s(u(t, x(t))) = s(v(t)).$$

The problem is now reduced to an ODE:

$$v' = s(v).$$

Going back to u:

$$u(t,x) = \int_0^t s(u(\tau, x(\tau)))\, d\tau = \tag{3.26}$$

$$= \int_0^t s(u(\tau, x - \lambda\tau))\, d\tau. \tag{3.27}$$

A similar procedure can be applied to linear constant coefficient systems: with notations as in (3.19), let $w = R^{-1}u$. Then we have:

$$w_t + \Lambda w_x = S(w),$$

where $S(w) = R^{-1}s(Rw)$. For every $p = 1, \ldots, m$ we can solve

$$(w_p)_t + \lambda_p (w_p)_x = S_p(w)$$

as in (3.26). Then we obtain our solution going back to u by $Rw = u$.

CHAPTER 4

Numerical methods for Balance Laws

1. Conservation Laws

We begin the review of numerical methods for conservation laws starting from the basic theory for the linear advection equation and linear constant coefficient hyperbolic systems.

We will usually consider a uniform grid in space $x_j = jh$, $h = \Delta x$, because it will be a necessary requirement for the methods we will use later, although most of the methods discussed can be extended to variable meshes. It will be useful to define also

$$x_{j+\frac{1}{2}} = x_j + \frac{h}{2}$$

and u_j^n as the computed approximation of the true solution $u(t_n, x_j)$. In finite difference schemes, u_j^n is the approximation of the point-wise value $u(t_n, x_j)$ while in finite volumes it is the approximation of the cell average of $u(t_n, x)$ in the cell $I_j = \left(x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}\right)$:

$$\bar{u}_j^n = \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} u(t_n, x)\, dx. \tag{4.1}$$

This interpretation is natural since the integral form of the conservation law describes precisely the time evolution of integrals such as that appearing in (4.1). We use $u_0(x)$ to define u_j^0 as initial data for the numerical method, either by pointwise values, $u_j^0 = u_0(x_j)$, or preferably by cell averages, $u_j^0 = \bar{u}_0(x_j)$.

In practice we compute the solution on a finite spatial domain, say $a \le x \le b$, so also appropriate boundary conditions have to be assigned to solve the problem completely. As usual, periodic conditions can be defined

$$u(t, a) = u(t, b) \qquad \forall t$$

but this is not always possible and boundary values have to be assigned according to the specific problem.

Numerical methods for hyperbolic CLs have been studied for a long time; a wide variety of both finite difference and finite volumes methods have been developed. Many of these are derived simply by replacing the derivatives occurring in (3.19) by appropriate finite difference approximations. For example, replacing u_t by a forward-in-time approximation and u_x by a spatially centered approximation, we obtain the following difference equations for u^{n+1}:

$$\frac{u_j^{n+1} - u_j^n}{k} + A\left(\frac{u_{j+1}^n - u_{j-1}^n}{2h}\right) = 0, \tag{4.2}$$

where $k = \Delta t$. This can be solved for u_j^{n+1} to obtain

$$u_j^{n+1} = u_j^n - \frac{k}{2h} A \left(u_{j+1}^n - u_{j-1}^n \right). \tag{4.3}$$

Unfortunately, despite the quite natural derivation of this method, it suffers from severe stability problems and is useless in practice.

Many other methods, having better properties, are defined in a similar fashion, using different grid points in space and/or in time.

If we look at the grid points involved in the computation of u_j^{n+1} with a given method, we obtain a diagram that is known as the *stencil* of the method.

A wide variety of methods can be devised for linear system (3.19) by using different finite difference approximations. A few possibilities are listed in Table 1. Most of them are based directly on finite difference approximations to the PDEs; an exception is the Lax-Wendroff method, which is based on the Taylor series expansion.

Name	Difference equation
Backward Euler	$u_j^{n+1} = u_j^n - \frac{k}{2h} A \left(u_{j+1}^n - u_{j-1}^n \right)$
One sided	$u_j^{n+1} = u_j^n - \frac{k}{h} A \left(u_j^n - u_{j-1}^n \right)$
One sided	$u_j^{n+1} = u_j^n - \frac{k}{h} A \left(u_{j+1}^n - u_j^n \right)$
Lax-Friedrichs	$u_j^{n+1} = \frac{1}{2} \left(u_{j+1}^n + u_{j-1}^n \right) - \frac{k}{2h} A \left(u_{j+1}^n - u_{j-1}^n \right)$
Leapfrog	$u_j^{n+1} = u_j^{n-1} - \frac{k}{2h} A \left(u_{j+1}^n - u_{j-1}^n \right)$
Lax-Wendroff	$u_j^{n+1} = u_j^n - \frac{k}{2h} A \left(u_{j+1}^n - u_{j-1}^n \right)$ $+ \frac{k^2}{2h^2} A^2 \left(u_{j+1}^n - 2u_j^n + u_{j-1}^n \right)$
Beam-Warming	$u_j^{n+1} = u_j^n - \frac{k}{2h} A \left(3u_j^n - 4u_{j-1}^n + u_{j-2}^n \right) +$ $+ \frac{k^2}{2h^2} A^2 \left(u_j^n - 2u_{j-1}^n + u_{j-2}^n \right)$

TABLE 1. Finite difference methods for the linear problem $u_t + Au_x = 0$.

We observe that all methods listed in Table 1 are *linear methods*: if we write

$$u^{n+1} = \mathcal{H}(u^n)$$

we have that $\mathcal{H}$ is a linear operator: this means equality

$$\mathcal{H}(\alpha u^n + \beta v^n) = \alpha \mathcal{H}(u^n) + \beta \mathcal{H}(v^n)$$

holds for grid functions u^n and v^n and scalar constants α and β. Linearity of difference methods is heavily used in the study of discrete approximations to linear PDEs: error analysis can be performed and convergence and stability results can be obtained for linear schemes. For a complete discussion about this argument, we refer to [**42**]; here we will state only the main results.

For linear difference schemes, a fundamental convergence theorem was given by Lax. Before we state the theorem, we recall the definition of *consistent* method for arbitrary *2-level* methods (with the exception of the Leapfrog method, all of the methods in Table 1 are 2-level methods; for time-dependent conservation laws, 2-level methods are almost exclusively used since those involving more than 2 levels have additional difficulties): let

$$L_k(t,x) = \frac{1}{k} \big[u(t+k,x) - \mathcal{H}_k(u(t,\cdot);x) \big] \tag{4.4}$$

be the *Local Truncation Error*, where $\mathcal{H}_k(u(t,\cdot);x)$ is the "continuous operator" (see [**42**]), then the method is *consistent* if

$$|L_k(t,x)| \to 0 \quad \text{as} \quad k \to 0. \tag{4.5}$$

Lax Equivalence Theorem says that for for a consistent, linear method, stability is necessary and sufficient for convergence.

A full proof may be found in [**52**].

1.1. The CFL condition. One of the first papers on finite difference methods for PDEs was written in 1928 by Courant, Friedrichs and Lewy ([**20**]). They used finite difference methods as an analytic tool for proving existence of solutions of certain PDEs. The idea is to first define a sequence of approximate solutions (via finite difference equations), then to prove that they converge as the grid is refined and finally show that the limit function must satisfy the PDE, giving the existence of a solution.

In the course of proving convergence of this sequence, they recognized that a necessary stability condition, not sufficient, for any numerical method is that the domain of dependence of the finite difference method should include the domain of dependence of the PDE, at least in the limit as the grid is refined. It is necessary since if there are points in the true domain of dependence that are not in the numerical domain of dependence, changing the value of the initial data at those points would thus effect the true solution but not the numerical solution, and hence the numerical solution cannot possibly converge to the true solution for all initial data.

This condition is known as the *CFL condition* after Courant, Friedrichs, and Lewy.

1.2. Upwind methods. For the scalar advection equation with $\lambda > 0$, the first of the two one-sided method in Table 1 can be applied and is stable provided the suitable CFL condition is satisfied. This method is usually called the first order *upwind method*, since the one-sided stencil points in the "upwind" or "upstream" direction, the correct direction from which characteristic information propagates. If we think of the advection equation as modeling the advection of a concentration profile in a fluid stream, then this is literally the upstream direction.

Similarly, the second of the two one-sided method in Table 1 is the upwind method for the advection equation with $\lambda < 0$. For a system of equations, we have seen that a one-sided method can only be used if all of the eigenvalues of A have the same sign. This is typically not the case in practice; for example, linearized Euler equations have mixed sign eigenvalues.

When computing discontinuous solutions, upwind differencing turns out to be an important tool, even for indefinite systems with eigenvalues of mixed sign. The appropriate application of upwind methods requires some sort of decomposition into characteristic fields. For example, if we change variables and decouple the linear system diagonalizing the coefficient matrix, then each of the resulting scalar problems can be solved with an appropriate upwind method, using the point to the left when $\lambda_p > 0$ or to the right when $\lambda_p < 0$.

For nonlinear systems analogous splitting can be introduced in various ways to incorporate upwinding. Many of these methods require solving Riemann problems in order to accomplish the appropriate splitting between wave propagation to the left and right.

1.3. Conservative Methods. When we attempt to numerically solve nonlinear conservation laws, we run into difficulties not seen in the linear equation. Nonlinearity makes everything harder to analyze. For smooth solutions to nonlinear problems, the numerical method can be linearized and results about linear finite difference methods can be applied to obtain convergence results for nonlinear problems. A very general theorem of this form is due to Strang ([**60**]).

For nonlinear problems the following difficulties could arise:

i)**:** the method might be "nonlinearly unstable", i.e., unstable on the nonlinear problem even though linearized versions appear to be stable. Often oscillations will trigger nonlinear instabilities;

ii)**:** the method might converge to a function that is not a weak solution of our original equation (or that is the wrong weak solution, i.e., does not satisfy the entropy condition).

Luckily, there turns out to be a very simple and natural requirement we can impose on our numerical methods which will guarantee that we do not converge to non-solutions. This is the requirement that the method be in *conservation form*, which means it has the following form:

$$u_j^{n+1} = u_j^n - \frac{k}{h}[F(u_{j-p}^n, \ldots, u_{j+q}^n) - F(u_{j-p-1}^n, \ldots, u_{j+q-1}^n)], \tag{4.6}$$

for some function F of $p+q+1$ arguments. F is called the *numerical flux function*. In the simplest case $p = 0$ and $q = 1$, so that F is a function of only two variables, (4.6) becomes

$$u_j^{n+1} = u_j^n - \frac{k}{h}[F(u_j^n, u_{j+1}^n) - F(u_{j-1}^n, u_j^n)]. \tag{4.7}$$

This form is very natural if we consider u_j^n as an approximation to the cell average $\bar{u}_j^n$.

Mo rover, a method in conservation form is *consistent* with the original conservation law if the numerical flux function F reduces to the true flux f in the case of constant flow: if $u(t,x) \equiv \bar{u}$, then we expect

$$F(\bar{u}, \ldots, \bar{u}) = f(\bar{u}), \qquad \forall \bar{u} \in \mathbb{R}.$$

Some smoothness is also required, so that as the arguments of F approach some common value $\bar{u}$, the value of F approaches $f(\bar{u})$ smoothly. For consistency it suffices to have F a Lipschitz continuous function of each variable.

It is very important to use conservative methods when dealing with CLs since for such schemes it is possible to prove a *discrete conservation* formula: a discrete form of the integral form of the conservation law (the total quantity of a conserved variable in any region changes only due to flux through the boundaries). For a conservative consistent scheme, using the notation $U_k(t,x)$ for the piecewise constant function defined by u_j^n, it is easy to see that the following discrete equivalent of conservation holds:

$$\int_{J-\frac{1}{2}}^{K+\frac{1}{2}} U_k(t_n, x)\, dx = \int_{J-\frac{1}{2}}^{K+\frac{1}{2}} u(t_n, x)\, dx$$

for some indices J and K sufficiently far out so that u_0 is constant outside some finite interval. Discrete conservation means that any shock we compute must, in a sense, be in

the "correct" location. The solution, computed with a conservative method, might have a smeared out shock but it must, at least, be smeared about the correct location.

1.4. Lax-Wendroff Theorem. We would like to correctly approximate discontinuous weak solutions to the conservation law by using a conservative method. Lax and Wendroff proved in [**38**] that this is true, at least in the sense that if we converge to some function $u(t, x)$ as the grid is refined, both in space and time with $\Delta t/\Delta x$ fixed, then this function will in fact be a weak solution of the conservation law. This theorem does not guarantee, instead, that we do converge: for that we need some form of stability, and even then if there is more than one weak solution, it might be that one sequence of approximations will converge to one weak solution, while another sequence converges to a different weak solution. Nonetheless, this is a very powerful and important theorem, for it says that we can have confidence in solutions we compute.

There are many examples of conservative numerical methods that converge to weak solutions violating the entropy condition: consider Burgers' equation with data:

$$u_0(x) = \begin{cases} -1 & x < 0, \\ +1 & x > 0. \end{cases}$$

The entropy satisfying weak solution consists of a rarefaction wave, but the stationary discontinuity $u(t, x) = u_0(x)$ is also a weak solution. Rankine-Hugoniot condition with $s = 0$ is satisfied since $f(-1) = f(1)$ for Burgers' equation. There are very natural conservative methods that converge to this latter solution rather than to the physically correct rarefaction wave.

For some numerical methods, it is possible to show that this can never happen, and that any weak solution obtained by refining the grid must in fact satisfy the entropy condition. Provided that the convergence can be ensured in some way, it can be shown that the weak solution obtained in the limit satisfies an entropy condition for a suitable entropy pair by showing that a discrete version of the entropy inequality (3.15) holds. The proof mimics that of Lax-Wendroff theorem; we refer to [**42**] and references therein for in-depth examination.

In general, the above conditions are far from easy to test for individual schemes; however, there do exist classes of schemes which are known to possess this entropy-satisfying property.

1.5. Godunov's Method. Recall that one-sided methods cannot be used for systems of equations with eigenvalues of mixed sign. For a linear system of equations we previously obtained a natural generalization of the upwind method by diagonalizing the system. For nonlinear systems the matrix of eigenvectors is not constant, and this same approach does not work directly. We will study a generalization in which the local characteristic structure, now obtained by solving a Riemann problem rather than by diagonalizing the Jacobian matrix, is used to define a natural upwind method. This method was first proposed for gas dynamics calculations by Godunov ([**29**]).

There are many reasons for introducing methods based on the solution of Riemann problems. Both Lax-Friedrichs and the upwind method are only first order accurate on smooth data, and even the less dissipative upwind method gives unacceptably smeared shock profiles. One would like to correct these deficiencies by developing "high resolution" methods that are second order accurate in smooth regions and give much sharper discontinuities.

The first order Godunov method forms a basis for many of the high resolution generalizations that have been developed.

A straightforward generalization of the upwind method to non linear systems would give a method that may compute entropy violating solutions; Godunov suggested solving Riemann problems forward in time. For any nonlinear scalar conservation law the correct Godunov flux which computes entropy condition satisfying weak solutions is:

$$F(u_l, u_r) = \begin{cases} \min_{u_l \le u \le u_r} f(u) & \text{if } u_l \le u_r \\ \max_{u_r \le u \le u_l} f(u) & \text{if } u_l > u_r. \end{cases} \tag{4.8}$$

This also follows from a more general result due to Osher ([**47**]), who found a closed form expression for the entropy solution $u(t,x) = w(x/t)$ of a general nonconvex scalar Riemann problem with data u_l and u_r.

1.6. Approximate Riemann Solvers. Godunov's method, and higher order variations of the method, require the solution of Riemann problems at every cell boundary in each time step. Although in theory these Riemann problems can be solved, in practice doing so is too expensive, and typically requires some iteration for nonlinear equations.

Moreover, most of the structure of the resulting Riemann solver is not used in Godunov's method. The exact solution is averaged over each grid cell, introducing large numerical errors. This suggests that it is not worthwhile calculating the Riemann solutions exactly and that we may be able to obtain equally good numerical results with an *approximate Riemann solution* obtained by some less expensive means.

One of the most popular Riemann solvers currently in use is due to Roe ([**54**]). The idea is to determine the solution by solving a constant coefficient linear system of conservation laws instead of the original nonlinear system: one solves a modified conservation law with flux $\widehat{f}(u) = \widehat{A}u$. Of course the coefficient matrix used to define this linear system must depend on u_l and u_r. To determine $\widehat{A}(u_l, u_r)$ in a reasonable way, Roe suggested that the following conditions should be imposed on A:

i): $\widehat{A}(u_l, u_r)(u_r - u_l) = f(u_r) - f(u_l)$;
ii): $\widehat{A}(u_l, u_r)$ is diagonalizable with real eigenvalues;
iii): $\widehat{A}(u_l, u_r) \to f'(\bar{u})$ smoothly as $u_l, u_r \to \bar{u}$.

For special systems of equations it is possible to derive suitable $\widehat{A}$ matrices that are very efficient to use relative to the exact Riemann solution; Roe showed how to do this for the Euler equations in [**54**]. We refer to the work of Roe [**54**] and [**28**] for further details and to [**42**] for examples of many other Riemann solvers.

2. Nonlinear Stability

Lax-Wendroff Theorem presented before does not say anything about whether a method converges or not; it only states that if a sequence of approximations converges, then the limit is a weak solution. To guarantee convergence, we need some form of stability, just as for linear problems. Unfortunately, the Lax Equivalence Theorem no longer holds and we cannot use the same approach (which relies heavily on linearity) to prove convergence.

We consider one form of nonlinear stability that allows us to prove convergence results for a wide class of practical methods. So far, this approach has been completely successful only for scalar problems. For general systems of equations with arbitrary initial data no

numerical method has been proved to be stable or convergent, although convergence results have been obtained in some special cases.

Even if the scalar case has limited direct applicability to real-world problems, it has been carefully studied because most of the successful numerical methods for systems, like Euler equations, have been developed by first inventing good methods for the scalar case (where theory provides good guidance) and then extending them in a relatively straightforward way to systems of equations. The fact that we can prove they work well for scalar equations is no guarantee that they will work at all for systems, but in practice this approach has been very successful.

2.1. Total Variation Stability. One difficulty immediately presents itself when we deal with the convergence of a numerical method for conservation laws: the standard definition of global error cannot be used when the weak solution is not unique. In order to prove a convergence result, we must first define an appropriate notion of "stability". For nonlinear problems, the primary tool used to prove convergence is compactness. A natural setting for solutions to our problem is the L^1 space and a way to ensure convergence is that of choosing the compact set of *total variation bounded* functions.

We will say that a numerical method is *total variation stable*, or simply TV-stable, if all approximations u_k for $k < k_0$ lie in some fixed set of the form

$$K = \{u \in L_{1,T} : TV_T(u) \leq R,\ \mathrm{Supp}(u(t,\cdot)) \subset [-M, M],\ \forall t \in [0,T]\}$$

where TV_T is the total variation over $[0,T]$ and R and M may depend on the initial data u_0 and the flux function $f(u)$, but not on k ([**42**]).

The TV-stability requirement can be simplified considerably in the special case of functions generated by conservative numerical methods: consider a conservative method with a Lipschitz continuous numerical flux F and suppose that for each initial data u_0 there exist some $k_0, R > 0$ such that

$$TV(u^n) \leq R \qquad \forall n, k \text{ with } k < k_0, nk \leq T.$$

Then the method is TV-stable. This means that if the one-dimensional total variation at each time t^n is uniformly bounded (independently on n), then global uniform boundedness follows.

The *fundamental convergence theorem* states that a conservative, consistent and TV-stable method is convergent. See [**42**] for a general formulation and a proof of the theorem.

A variety of classes of methods have been proved to be TV-stable and hence will be convergent when in a conservative and consistent form. As examples, we recall *Total Variation Diminishing* and *monotonicity preserving* methods.

A numerical method is called Total Variation Diminishing (TVD) if

$$TV(u^{n+1}) \leq TV(u^n).$$

It can be shown that the true solution of a scalar conservation law has this TVD property: any weak solution $u(t,x)$ satisfies

$$TV(u(t_2,x)) \leq TV(u(t_1,x)) \qquad \forall t_2 \geq t_1.$$

If this were not the case then it would be impossible to develop a TVD numerical method; however, since true solutions are TVD, it is reasonable to impose this requirement on the numerical solution as well, yielding a TV-stable and hence convergent method. A number

of very successful numerical methods have been developed using this requirement, among which we recall Runge-Kutta type TVD schemes.

2.2. Monotonicity Preserving methods. Recall that one difficulty associated with numerical approximations of discontinuous solutions is that oscillations may appear near the discontinuity. In an attempt to eliminate this unwanted defect, one natural requirement we might place on a numerical method is that it be *monotonicity preserving*. This means that if the initial data is monotone as a function of space, then the solution should have the same property for all times n. This means in particular that oscillations cannot arise near an isolated propagating discontinuity, since the Riemann initial data is monotone. An important result is that any TVD method is monotonicity preserving ([**42**]).

Another attractive feature of the TVD requirement is that it makes deriving methods with a high order of accuracy which are TVD possible. By contrast, if we defined "stability" by mimicking certain other properties of the true solution, we would find that accuracy is limited to first order.

3. High resolution schemes for homogeneous conservation laws

Monotone methods for scalar conservation laws are TVD and satisfy a discrete entropy condition. Hence they converge in a nonoscillatory manner to the unique entropy solution. However, monotone methods are at most first order accurate, giving poor accuracy in smooth regions of the flow. Moreover, shocks tend to be heavily smeared and poorly resolved on the grid. These effects are due to the large amount of numerical dissipation in monotone methods. Some dissipation is obviously needed to give nonoscillatory shocks and to ensure that we converge to the vanishing viscosity solution, but monotone methods go overboard in this direction.

To overcome these difficulties, *high resolution* methods have been developed and studied. This term applies to methods that are at least second order accurate on smooth solutions and yet give well resolved, nonoscillatory discontinuities.

In the scalar problem, the constraint that the method be total variation diminishing can be imposed. This insures/ensures that we obtain nonoscillatory shocks and convergence in the sense of Lax-Wendroff Theorem. These scalar methods will later be extended to systems of equations using an approximate decomposition of the system into characteristic fields.

The main idea behind any high resolution method is to attempt to use a high order method, but to modify the method and increase the amount of numerical dissipation in the neighborhood of a discontinuity.

There exists a wide variety of available approaches, and often there are close connections between the methods developed by quite different means. We will name here just three classes of quite popular methods:

- artificial dissipation;
- flux-limiting;
- slope-limiting

and refer to the literature for details ([**28, 42**]).

3.1. Semi-discrete Methods. Methods discussed so far have all been fully discrete methods, i.e. discretizing in both space and time. At times it is useful to consider the

discretization process at two different stages, first discretizing only in space, leaving the problem continuous in time. This leads to a system of ordinary differential equations in time, called the *semi-discrete equations*. We then discretize in time using any standard numerical method for systems of ordinary differential equations. This approach of reducing a PDE to a system of ODEs, to which we then apply an ODE solver, is often called the *method of lines*.

This approach is particularly useful in developing methods with order of accuracy greater than 2, since it allows us to decouple the issues of spatial and temporal accuracy. We can define high order approximations of the flux at a cell boundary at one instant in time using high order interpolation in space, and then achieve high order temporal accuracy by applying any of the wide variety of high order ODE solvers. This approach is also useful in extending methods to two or more space dimensions.

Many high order semi-discrete methods have been developed; we refer to [**42, 28**] for a general overview. Here we will only recall the basic theory about *Essentially Non-Oscillatory* (ENO) and *Weighted Essentially Non-Oscillatory* (WENO) schemes. ENO and WENO are high order accurate finite volumes or finite difference schemes designed for problems with piecewise smooth solutions containing discontinuities.

The key idea lies at the approximation level, where a nonlinear adaptive procedure is used to automatically choose the locally smoothest stencil, hence avoiding crossing discontinuities in the interpolation procedure as much as possible. ENO and WENO schemes have been quite successful in applications, especially for problems containing both shocks and complicated smooth solution structures (such as compressible turbulence simulations and aeroacoustics).

Since the publication of the original paper of Harten, Engquist, Osher and Chakravarthy ([**33**]), many researchers have studied this pioneer work, improving the methodology and expanding the area of its applications. ENO schemes based on point values and TVD Runge-Kutta time discretizations, which can save computational costs significantly for multi space dimensions, were developed in [**58**] and [**59**]. Weighted ENO schemes were then developed, using a convex combination of all candidate stencils instead of just one as in the original ENO ([**35, 44**]). We refer to [**57**] as a comprehensive reference about ENO and WENO schemes; the material that follows is also inspired by this last article.

The first approximation problem we will face, in solving hyperbolic conservation laws using cell averages is the *reconstruction* problem. We will assume for the moment a non uniform grid.

Reconstruction means the following: given the cell averages of a function $v(x)$:

$$\bar{v}_i \equiv \frac{1}{\Delta x_i} \int_{x_i-\frac{1}{2}}^{x_i+\frac{1}{2}} v(\xi)\, d\xi, \qquad i = 1, \ldots, N \tag{4.9}$$

find a polynomial $p_i(x)$, of degree at most $k-1$, for each cell I_i, such that it is a k-th order accurate approximation to the function $v(x)$ inside I_i:

$$p_i(x) = v(x) + O(\Delta x^k), \qquad x \in I_i, \quad i = 1, \ldots, N, \tag{4.10}$$

where $\Delta x = max_i \Delta x_i$. In particular, this gives approximations to the function $v(x)$ at the cell boundaries

$$v^-_{i+\frac{1}{2}} = p_i(x_{i+\frac{1}{2}}), \quad v^+_{i-\frac{1}{2}} = p_i(x_{i-\frac{1}{2}}), \quad i = 1, \ldots, N, \tag{4.11}$$

which are $k-$th order accurate:

$$v^-_{i+\frac{1}{2}} = v(x_{i+\frac{1}{2}}) + O(\Delta x^k), \quad v^+_{i-\frac{1}{2}} = v(x_{i-\frac{1}{2}}) + O(\Delta x^k), \quad i = 1, \ldots, N, \tag{4.12}$$

Polynomials p_i can be replaced by other functions. We will not discuss boundary conditions: we assume that $\bar{v}_i$ is also available for $i < 0$ and $i > N$ if needed.

Given the location I_i and the order of accuracy k, we first choose a stencil, based on r cells to the left, s cells to the right, and I_i itself; if $r, s \geq 0$, with $r + s + 1 = k$,

$$S(i) = \{I_{i-r}, \ldots, I_{i+s}\}. \tag{4.13}$$

There is a unique polynomial of degree at most $k - 1 = r + s$, denoted by $p(x)$, whose cell average in each of the cells in $S(i)$ agrees with that of $v(x)$; such polynomial $p(x)$ is the $k-$th order approximation we are looking for. To solve our problem we also need the approximations to the values of $v(x)$ at the cell boundaries. Since the mappings from the given cell averages $\bar{v}_j$ in the stencil $S(i)$ to the values $v^-_{i+\frac{1}{2}}$ and $v^+_{i-\frac{1}{2}}$ in (4.11) are linear, there exist constants c_{rj} and $\widetilde{c}_{rj}$, which depend on the left shift r of the stencil $S(i)$, on the order of accuracy k, and on the cell sizes Δx_j in the stencil $S(i)$, but not on the function v itself, such that

$$v^-_{i+\frac{1}{2}} = \sum_{j=0}^{k-1} c_{rj} \bar{v}_{i-r+j}, \qquad v^+_{i-\frac{1}{2}} = \sum_{j=0}^{k-1} \widetilde{c}_{rj} \bar{v}_{i-r+j}. \tag{4.14}$$

In fact, it can be shown that $\widetilde{c}_{rj} = c_{r-1\,j}$, so it is possible to drop the superscripts $\pm$ in (4.14).

In summary, given the k cell averages of the function v in the stencil $S(i)$, there are constants c_{rj} such that the reconstructed value at the cell boundary $x_{i+\frac{1}{2}}$

$$v_{i-\frac{1}{2}} = \sum_{j=0}^{k-1} \widetilde{c}_{rj} \bar{v}_{i-r+j} \tag{4.15}$$

is $k-$order accurate:

$$v_{i+\frac{1}{2}} = v(x_{i+\frac{1}{2}}) + O(\Delta x^k), \quad . \tag{4.16}$$

We list, in Table 2, constants c_{rj} for the uniform grid case, for order of accuracy between $k = 1$ and $k = 3$ (refer to [**57**] higher orders of accuracy and for the general non uniform grid case). The second approximation problem is obtaining high order *conservative* approximation to the derivative from point values: given the point values of a function $v(x)$

$$v_i \equiv v(x_i), \qquad i = 1, \ldots, N$$

find a numerical flux function

$$\widehat{v}_{i+\frac{1}{2}} \equiv \widehat{v}(v_{i-r}, \ldots, v_{i+s}), \qquad i = 1, \ldots, N \tag{4.17}$$

k	r	$j=0$	$j=1$	$j=2$
1	-1	1		
	0	1		
2	-1	3/2	-1/2	
	0	1/2	1/2	
	1	-1/2	3/2	
3	-1	11/6	-7/6	1/3
	0	1/3	5/6	-1/6
	1	-1/6	5/6	1/3
	2	1/3	-7/6	11/6

TABLE 2. Constants c_{rj} for the uniform grid case

such that the flux difference approximates the derivative $v'(x)$ to $k-$th order accuracy:

$$\frac{1}{\Delta x_i}\left(\widehat{v}_{i+\frac{1}{2}} - \widehat{v}_{i-\frac{1}{2}}\right) = v'(x_i) + O(\Delta x^k), \qquad i = 1, \ldots, N. \tag{4.18}$$

We assume that the grid is uniform, $\Delta x_i = \Delta x$. This assumption is, unfortunately, essential in the following development: otherwise, it can be proven that *no* choice of constants c_{rj} could make the conservative approximation to the derivative higher than second order accurate ($k > 2$). Thus, the high order (third order and higher) finite difference schemes discussed here can apply only to uniform or smoothly varying grids.

This problem looks quite different from the previous one but, instead, they are strictly related: if we can find a function $h(x)$, which may depend on the grid size Δx, such that

$$v(x) = \frac{1}{\Delta x}\int_{x-\frac{\Delta x}{2}}^{x+\frac{\Delta x}{2}} h(\xi)\, d\xi, \tag{4.19}$$

then clearly

$$v'(x) = \frac{1}{\Delta x}\left[h\left(x + \frac{\Delta x}{2}\right) - h\left(x - \frac{\Delta x}{2}\right)\right].$$

The known function $v(x)$ is the cell average of the unknown function $h(x)$, so to find $h(x)$ we just need to use the reconstruction procedure described before. Thus, given the point values $\{v_j\}$, we "identify" them as cell averages of another function $h(x)$ and so a primitive function of $h(x)$ is exactly known at the cell interfaces $x = x_{i+\frac{1}{2}}$: we thus use the same reconstruction procedure described before to get a k-th order approximation to $h(x_{i+\frac{1}{2}})$, which is then taken as the numerical flux $\widehat{v}_{i+\frac{1}{2}}$ in (4.17).

3.2. Essentially Non Oscillatory scheme. The heart of ENO scheme is the idea of *adaptive stencil*: namely, the left shift r changes with the location x_i. The basic idea is to avoid including the discontinuous cell in the stencil, if possible.

Let $V[x_{i-\frac{1}{2}}] \equiv V(x_{i-\frac{1}{2}})$ be the 0-th degree divided differences of the function $V(x)$ and the j-th degree divided differences, for $j \geq 1$, be defined inductively by

$$V[x_{i-\frac{1}{2}}, \ldots, x_{i+j-\frac{1}{2}}] \equiv \frac{V[x_{i+\frac{1}{2}}, \ldots, x_{i+j-\frac{1}{2}}] - V[x_{i-\frac{1}{2}}, \ldots, x_{i+j-\frac{3}{2}}]}{x_{i+j-\frac{1}{2}} - x_{i-\frac{1}{2}}}. \tag{4.20}$$

We recall a fundamental property of divided differences:

$$V[x_{i-\frac{1}{2}},\ldots,x_{i+j-\frac{1}{2}}]=\frac{V^{(j)}(\xi)}{j!}$$

for some ξ inside the stencil as long as the function $V(x)$ is smooth in this stencil. If $V(x)$ is discontinuous at some point inside the stencil, then it is easy to verify that

$$V[x_{i-\frac{1}{2}},\ldots,x_{i+j-\frac{1}{2}}]=O(\frac{1}{\Delta x^j})$$

Thus the divided difference is a measurement of the smoothness of the function inside the stencil.

The ENO idea is the following: suppose we want to find a stencil of $k+1$ consecutive points, which must include $x_{i-\frac{1}{2}}$ and $x_{i+\frac{1}{2}}$, such that $V(x)$ is *the smoothest* in this stencil comparing with other possible stencils. This can be done by breaking the procedure into steps, where in each step we only add one point to the stencil. We thus start with the two point stencil; at the next step, we have only two choices to expand the stencil by adding one point: we can either add the left or the right neighbor. We have already noticed before, that a smaller divided difference implies the function is smoother in that stencil. We thus decide upon which point to add to the stencil by comparing the two relevant divided differences, and picking the one with a smaller absolute value. Thus, if

$$|V[x_{i-\frac{3}{2}},x_{i-\frac{1}{2}},x_{i+\frac{1}{2}}]|<|V[x_{i-\frac{1}{2}},x_{i+\frac{1}{2}},x_{i+\frac{3}{2}}]|$$

we will take the 3 point stencil as $\{x_{i-\frac{3}{2}},x_{i-\frac{1}{2}},x_{i+\frac{1}{2}}\}$ otherwise, we will choose $\{x_{i-\frac{1}{2}},x_{i+\frac{1}{2}},x_{i+\frac{3}{2}}\}$.

This procedure can be continued, with one point added to the stencil at each step, according to the smaller of the absolute values of the two relevant divided differences, until the desired number of points in the stencil is reached. Once the stencil is found, one could use (4.15) to compute the reconstructed values at the cell boundary or one could use it to compute the fluxes. An alternative way is to compute the values or fluxes using the Newton form directly ([**57**]).

For a piecewise smooth function $v(x)$, ENO interpolation starting with a two point stencil has the following properties:

- The accuracy condition

  $$p_i(x)=v(x)+O(\Delta x^{k+1}),\qquad x\in I_i$$

  is valid for any cell I_i which does not contain a discontinuity. This implies that the ENO interpolation procedure can recover the full high order accuracy right up to the discontinuity;
- $p_i(x)$ is monotone in any cell I_i which does contain a discontinuity of $V(x)$;
- The reconstruction is TVB (total variation bounded).

3.3. Weighted Essentially Non Oscillatory scheme. ENO reconstruction is uniformly high order accurate right up to the discontinuity. It achieves this effect by adaptively choosing the stencil based on the absolute values of divided differences. However, one could make the following remarks about ENO reconstruction, indicating rooms for improvements:

- the stencil might change even by a round-off error perturbation near zeros of the solution and its derivatives. This may cause loss of accuracy when applied to a hyperbolic PDE;
- the resulting numerical flux is not smooth, as the stencil pattern may change at neighboring points;
- only one of the stencils is actually used in forming the reconstruction while if all the $2k-1$ cells in the potential stencils were used, one could get higher order accuracy in smooth regions;
- ENO stencil choosing procedure involves many logical "if" structures which are not very efficient.

WENO is an improvement upon ENO. The basic idea is the following: instead of using only one of the candidate stencils to form the reconstruction, one uses a convex combination of all of them. To be more precise, suppose the k candidate stencils

$$S_r(i) = \{x_{i-r}, \ldots, x_{i-r+k-1}\}, \qquad r = 0, \ldots, k-1 \tag{4.21}$$

produce k different reconstructions to the value $v_{i+\frac{1}{2}}$:

$$v_{i+\frac{1}{2}}^{(r)} = \sum_{j=0}^{k-1} c_{rj} \bar{v}_{i-r+j}, \qquad r = 0, \ldots, k-1. \tag{4.22}$$

WENO reconstruction would take a convex combination of all $v_{i+\frac{1}{2}}^{(r)}$ as a new approximation to the cell boundary value

$$v_{i+\frac{1}{2}} = \sum_{r=0}^{k-1} \omega_r v_{i+\frac{1}{2}}^{(r)}. \tag{4.23}$$

Regarding weights ω_r we require

$$\omega_r \geq 0, \qquad \sum_{r=0}^{k-1} \omega_r = 1$$

for stability and consistency.

If $v(x)$ is smooth in all of the candidate stencils, then there are constants d_r such that

$$v_{i+\frac{1}{2}} = \sum_{r=0}^{k-1} d_r v_{i+\frac{1}{2}}^{(r)} = v(x_{i+\frac{1}{2}}) + O(\Delta x^{2k-1}). \tag{4.24}$$

For example, for $1 \leq k \leq 3$ d_r are given by

$$\begin{array}{lll} k=1, & d_0 = 1; & \\ k=2, & d_0 = \frac{2}{3}, & d_1 = \frac{1}{3}; \\ k=3, & d_0 = \frac{3}{10}, & d_1 = \frac{6}{10} \quad d_2 = \frac{1}{10}. \end{array}$$

In this smooth case, we would like to have

$$\omega_r = d_r + O(\Delta x^{k-1}) \qquad r = 0, \ldots, k-1$$

so that

$$v_{i+\frac{1}{2}} = \sum_{r=0}^{k-1} \omega_r v_{i+\frac{1}{2}}^{(r)} = v(x_{i+\frac{1}{2}}) + O(\Delta x^{2k-1}).$$

When the function $v(x)$ has a discontinuity in one or more of the stencils, we would hope the corresponding weights ω_r to be essentially 0, to emulate the successful ENO idea. Moreover, the weights should be smooth functions of the cell averages involved (and in fact, this will be the case) and also computationally efficient.

All these considerations lead to the following form of weights:

$$\omega_r = \frac{\alpha_r}{\sum_{i=0}^{k-1} \alpha_i}, \qquad r = 0, \ldots, k-1 \tag{4.25}$$

with

$$\alpha_r = \frac{d_r}{(\varepsilon + \beta_r)^2}. \tag{4.26}$$

Here $\varepsilon > 0$ is introduced to avoid the denominator to become 0 and β_r are the so-called *smoothness indicators* of the stencil $S_r(i)$. A very efficient derivation of smoothness indicators can be found in [**35**]: when $k = 2$

$$\begin{aligned} \beta_0 &= (\bar{v}_{i+1} - \bar{v}_i)^2, \\ \beta_1 &= (\bar{v}_i - \bar{v}_{i-1})^2, \end{aligned} \tag{4.27}$$

while for $k = 3$

$$\begin{aligned} \beta_0 &= \frac{13}{12}(\bar{v}_i - 2\bar{v}_{i+1} + \bar{v}_{i+2})^2 + \frac{1}{4}(3\bar{v}_i - 4\bar{v}_{i+1} + \bar{v}_{i+2})^2, \\ \beta_1 &= \frac{13}{12}(\bar{v}_{i-1} - 2\bar{v}_i + \bar{v}_{i+1})^2 + \frac{1}{4}(\bar{v}_{i-1} - \bar{v}_{i+1})^2, \\ \beta_2 &= \frac{13}{12}(\bar{v}_{i-2} - 2\bar{v}_{i-1} + \bar{v}_i)^2 + \frac{1}{4}(\bar{v}_{i-2} - 4\bar{v}_{i-1} + 3\bar{v}_i)^2. \end{aligned} \tag{4.28}$$

With this choice of smoothness indicators, (4.27) gives a third order WENO scheme and (4.28) a fifth order one.

Notice that the scheme discussed here has a one point upwind bias in the optimal linear stencil, suitable for a problem with wind blowing from left to right. If the wind blows the other way, the procedure should be modified symmetrically with respect to $x_{i+\frac{1}{2}}$.

In summary: given the cell averages $\bar{v}_i$ of a function $v(x)$, for each cell I_i, we obtain upwind biased $(2k-1)-$th order approximations to the function $v(x)$ at the cell boundaries, denoted by $v_{i-\frac{1}{2}}^+$ and $v_{i+\frac{1}{2}}^-$, in the following way:

(1) compute the k reconstructed values $v_{i+\frac{1}{2}}^{(r)}$ as in (4.22) and the k reconstructed values $v_{i-\frac{1}{2}}^{(r)}$ as in (4.14), based on the stencils (4.13) for $r = 0, \ldots, k-1$;

(2) find constants d_r and $\widetilde{d}_r$ such that (4.24) and

$$v_{i-\frac{1}{2}} = \sum_{r=0}^{k-1} \widetilde{d}_r v_{i-\frac{1}{2}}^{(r)} = v(x_{i-\frac{1}{2}}) + O(\Delta x^{2k-1})$$

are valid; by symmetry $\widetilde{d}_r = d_{k-1-r}$;

(3) find smoothness indicators β_r for $r = 0, \ldots, k-1$. Explicit formulae for $k = 2$ and $k = 3$ are given in (4.27) and (4.28) respectively;
(4) form weights ω_r and $\widetilde{\omega}_r$ using (4.25) - (4.26) and

$$\widetilde{\omega}_r = \frac{\widetilde{\alpha}_r}{\sum_{r=0} k-1\widetilde{\alpha}_r}, \qquad \widetilde{\alpha}_r = \frac{\widetilde{d}_r}{(\varepsilon + \beta_r)^2}, \qquad r = 0, \ldots, k-1;$$

(5) compute the $(2k-1)-$th order reconstructions:

$$v^-_{i+\frac{1}{2}} = \sum_{j=0}^{k-1} \omega_r v^{(r)}_{i+\frac{1}{2}}, \qquad v^+_{i-\frac{1}{2}} = \sum_{j=0}^{k-1} \widetilde{\omega}_r v^{(r)}_{i-\frac{1}{2}}.$$

We recall now the standard ENO and WENO procedures for 1D conservation laws:

$$u_t(t,x) + f_x(u(t,x)) = 0 \tag{4.29}$$

assuming suitable boundary conditions are given.

Finite volumes 1D scalar ENO and WENO formulation:

- find the $k-$th order reconstructed values $u^+_{i-\frac{1}{2}}$ and $u^-_{i+\frac{1}{2}}$ for all i, using ENO or WENO schemes described above;
- choose a monotone flux ([**57**]) to compute the flux $\widehat{f}_{i+\frac{1}{2}}$ for all i;
- form the scheme:

$$\frac{d\bar{u}_i}{dt}(t) = -\frac{1}{\Delta x}\left(\widehat{f}_{i+\frac{1}{2}} - \widehat{f}_{i-\frac{1}{2}}\right).$$

For the finite difference formulation, we solve (4.29) directly using a conservative approximation to the spatial derivative:

$$\frac{du_i}{dt}(t) = -\frac{1}{\Delta x}\left(\widehat{f}_{i+\frac{1}{2}} - \widehat{f}_{i-\frac{1}{2}}\right) \tag{4.30}$$

where $u_i(t)$ is the numerical approximation to the point value $u(t, x_i)$, and the numerical flux $\widehat{f}_{i+\frac{1}{2}}$ is obtained by the ENO or WENO reconstruction procedures, with $\bar{v}(x) = f(u(t,x))$.

For stability, it is important that upwinding is used in constructing the flux. The easiest and the most inexpensive way to achieve upwinding to choose ENO-Roe or WENO-Roe scheme ([**54**]) but in many cases such method in known to produce entropy violating solutions.

It is usually more robust to use *global flux splitting*:

$$f(u) = f^+(u) + f^-(u) \tag{4.31}$$

where

$$\frac{df^+}{du}(u) \geq 0, \qquad \frac{df^-}{du}(u) \leq 0. \tag{4.32}$$

We would need the positive and negative fluxes $f^{\pm}(u)$ to have as many derivatives as the order of the scheme. Examples of global flux-splitting can be found in [**57**].

Finite difference 1D scalar ENO and WENO formulation using flux splitting:

(1) find a smooth flux splitting (4.31), satisfying (4.32);
(2) identify $\bar{v}_i = f^+(u_i)$ and use the ENO or WENO reconstruction procedures to obtain the cell boundary values $v^-_{i+\frac{1}{2}}$ for all i;

(3) take the positive numerical flux as

$$\widehat{f}^{+}_{i+\frac{1}{2}} = v^{-}_{i+\frac{1}{2}};$$

(4) identify $\bar{v}_i = f^{-}(u_i)$ and use the ENO or WENO reconstruction procedures to obtain the cell boundary values $v^{+}_{i+\frac{1}{2}}$ for all i;

(5) take the negative numerical flux as

$$\widehat{f}^{-}_{i+\frac{1}{2}} = v^{+}_{i+\frac{1}{2}};$$

(6) form scheme (4.30).

Notice that the schemes are nonlinear also for linear constant coefficients PDEs.

Regarding the treatment of boundary conditions we resume here only the standard ideas. For periodic boundary conditions there is no difficulty: one simply set as many ghost points as needed using either the periodicity condition or the compactness of the solution. Other types of boundary conditions should be handled according to their type. For inflow boundary conditions, one would usually use the physical inflow boundary condition at the exact boundary; the same holds for outflow. Apart from that, the most natural way of treating boundary conditions for both ENO and WENO schemes is to use only the available values inside the computational domain when choosing the stencil: only stencils completely contained inside the computational domain should be used. In practical implementation, one could set all the ghost values outside the computational domain to be very large with large variations, e.g. setting $u_j = 10^{-j}$ at ghost points x_j: this way the procedure will automatically avoid choosing any stencil containing ghost points for ENO schemes or assign zero weight in WENO schemes. Another way of treating boundary conditions is to use extrapolation of suitable order to set the values of the solution in all necessary ghost points.

There are several ways to generalize scalar ENO or WENO schemes to systems.

The easiest way is to apply the ENO or WENO schemes in a component by component fashion, solving the problem for each component using the finite volume or the finite difference formulation described before. These component by component versions of ENO and WENO schemes are simple and cost effective and work reasonably well for many problems. However, for more demanding test problems, we would need the more costly, but much more robust characteristic decompositions.

The linear constant coefficient case $f(u) = A\,u$ can be treated in the "usual" way, switching to characteristic variables to obtain a system of decoupled linear advection equations. We can then use the reconstruction or flux evaluation techniques for the scalar equations to handle each of the equations. After we can go back to the "physical space".

In the general nonlinear case, where $f'(u)$ is not constant, the problem is that all matrices $R(u)$, $R^{-1}(u)$ and $\Lambda(u)$ (as in (3.20)) are dependent upon u. We must "freeze" them locally in order to carry out a similar procedure as in the constant coefficient case. Thus, to compute the flux at the cell boundary $x_{i+\frac{1}{2}}$, we would need an approximation to the Jocobian at the middle value $u_{i+\frac{1}{2}}$. This can be simply taken as the arithmetic mean

$$u_{i+\frac{1}{2}} = \frac{1}{2}(u_i + u_{i+1}) \tag{4.33}$$

or as a more elaborate average satisfying some nice properties, e.g. Roe average ([**54**]). Once we have $u_{i+\frac{1}{2}}$, we will use $R(u_{i+\frac{1}{2}})$, $R^{-1}(u_{i+\frac{1}{2}})$ and $\Lambda(u_{i+\frac{1}{2}})$ to evaluate the numerical flux. We then repeat the procedure described above for linear systems, the difference here being that the matrices are different at different locations, hence the cost of the operation is greatly increased.

Depending on which approximation scheme one chooses at each step, i.e. ENO or WENO in finite volume or Roe-type finite difference or flux-splitting finite difference form, the corresponding procedure for nonlinear system follows.

4. Multidimensional Problems

Most practical problems are in two or more space dimensions. So far, we have only considered the one-dimensional (1D) problem. To some extent the 1D methods and theory can be applied to problems in more than one space dimension, and some of these extensions will be briefly described here. We look at the two-dimensional (2D) case because it is the case of our interest and to keep notations simple, but the same ideas can be used in three dimensions as well.

In two space dimensions a system of conservation laws takes the form

$$u_t + f_x(u) + g_y(u) = 0 \tag{4.34}$$

where $u = u(t, x, y) \in \mathbb{R}^m$. Typically the problem geometry is complicated and this introduces great difficulties; since this is not needed here, we will only discuss the case in which a rectangular grid is used.

One approach to solve the 2D problem is to first introduce a semi-discrete approximation, as it was the case in 1D, and then discretize the resulting system of ODEs.

The simplest method of this form is a generalization of Godunov's method to 2D obtained solving two 1D Riemann problems, first in one direction and then using the resulting state as starting point for the other direction. Discretizing in time using Euler's method then gives a 2D generalization of Godunov's method, which is first order accurate. Higher order accuracy can be obtained using higher order approximations of the spatial derivatives and higher order accuracy in time as well.

Another approach to develop numerical methods in two space dimensions is to use any of the fully discrete one-dimensional methods presented before and to apply them alternately on one-dimensional problems in the x and y directions. In the scalar linear case, such technique solves the problem exactly; in particular, if we use two order p methods in both directions, the resulting method will maintain order p accuracy; this implies, also, that there is no splitting error. This is not true any more for systems, also in the constant coefficient case. Indeed, if we replaced the 2D solution operator by the product of two 1D solution operators (i.e. first we approximate in one direction and then in the other), we would introduce an error which depends on the commutator $AB - BA$, where $f'(u) = Au$ and $g'(u) = Bu$. If A and B do not commute, the splitting error degenerate the global method to first order accuracy. A way to obtain at least second order accuracy is using *Strang splitting* (presented in [**61**]).

It was shown in [**41**] that global second order accuracy is achieved also in the nonlinear case for smooth solutions; for nonsmooth solutions, it was shown ([**21**]) that convergence to the entropy satisfying weak solution can be ensured using monotone methods for the 1D

approximation (hence no better than first order accuracy). Unfortunately it was proved ([**30**]) that fully 2D TVD methods are at most first order accurate. In spite of this negative result, numerical methods obtained using high resolution 1D methods combined with the Strang splitting typically work very well in practice.

This is true also for the high resolution ENO and WENO schemes. For 2D problems they are commonly adopted in a method of line fashion:

$$\frac{\partial u_{ij}}{\partial t}(t) = -\frac{1}{\Delta x}\left(\widehat{f}_{i+\frac{1}{2},j} - \widehat{f}_{i-\frac{1}{2},j}\right) - \frac{1}{\Delta y}\left(\widehat{g}_{i,j+\frac{1}{2}} - \widehat{g}_{i,j-\frac{1}{2}}\right).$$

Both for accuracy and computational reasons, it is highly recommended to use the finite difference ENO or WENO schemes instead of the finite volumes formulations. Even though there are very little theoretical results about ENO or WENO schemes for the multidimensional case, in practice these schemes are very robust and stable.

The development of fully multidimensional methods (and the required mathematical theory!) is one of the exciting challenges for the future in this field.

5. Time evolution

Up to now we have only considered spatial discretizations, leaving the time variable continuous (method of lines). In this section we consider the issue of time discretization. We present a class of TVD high order Runge-Kutta methods, which were developed in [**58**] and further in [**31**].

These Runge-Kutta methods are used to solve systems of initial value problems of ODEs written as:

$$u_t = L(u), \tag{4.35}$$

resulting from a method of lines spatial approximation to a PDE such as

$$u_t = -f_x(u), \tag{4.36}$$

What we say here regarding 1D conservation laws apply also to general initial value problems of PDEs in any spatial dimensions. Clearly, $L(u)$ in (4.35) is an approximation (e.g. ENO or WENO approximation in our case), to the derivative $-f_x(u)$ in the PDE (4.36).

What is needed is a condition on the time step Δt, say $\Delta t \leq c$ (c is called the CFL coefficient), such that a high order Runge-Kutta method satisfies a stability condition

$$|u^{n+1}| \leq |u^n| \tag{4.37}$$

in a certain norm. Originally in [**58**] the norm in (4.37) was chosen to be the total variation norm, hence the terminology *TVD time discretization*.

For a general Runge-Kutta method

$$\begin{aligned} u^{(i)} &= \sum_{k=0}^{i-1}\left(\alpha_{ik}u^{(k)} + \Delta t\beta_{ik}L(u^{(k)})\right), \quad i = 1,\ldots,m \\ u^{(0)} &= u^n, \qquad u^{(m)} = u^{n+1}. \end{aligned} \tag{4.38}$$

the following result holds ([**58**]): if $\alpha_{ik}, \beta_{ik} \geq 0$, then the Runge-Kutta method (4.38) is TVD under the CFL coefficient

$$c = \min_{i,k}\frac{\alpha_{ik}}{\beta_{ik}}.$$

In [**31**], the optimal second and third order TVD Runge-Kutta methods are given, both resulting with a CFL coefficient $c = 1$: the optimal second order scheme reads

$$\begin{aligned} u^{(1)} &= u^{(n)} + \Delta t L(u^n), \\ u^{n+1} &= \tfrac{1}{2}u^n + \tfrac{1}{2}u^{(1)} + \tfrac{1}{2}\Delta t L(u^{(1)}) \end{aligned}$$

and the optimal third order TVD Runge-Kutta method is given by:

$$\begin{aligned} u^{(1)} &= u^{(n)} + \Delta t L(u^n), \\ u^{(2)} &= \tfrac{3}{4}u^n + \tfrac{1}{4}u^{(1)} + \tfrac{1}{4}\Delta t L(u^{(1)}) \\ u^{n+1} &= \tfrac{1}{3}u^n + \tfrac{2}{3}u^{(2)} + \tfrac{2}{3}\Delta t L(u^{(2)}). \end{aligned} \tag{4.39}$$

Also fourth and fifth order TVD-RK schemes are available; we refer to [**31**] for their explicit formulation and for other aspects related to the theory of TVD-RK schemes, such as multistep methods and storage issues.

6. Balance Laws

There are various ways to handle source terms, which fall into two basic categories:

- *unsplit* methods, in which a single finite-difference formula is developed to advance the full equation over one time step;
- *fractional step* (splitting) methods, in which the problem is divided into pieces corresponding to different processes and a numerical method appropriate for each separate piece is applied independently.

A classical example of fractional step (also called operator splitting) technique, is given by the alternate solution of the simpler problems

$$u_t + f_x(u) = 0$$

and

$$u_t = s(u).$$

This approach is quite simple and it allows to use high-resolution methods for conservation laws without change, coupling these methods with standard ODE solvers for the second problem: one can advance in time simply alternating one solver to the other. Although benefiting from simplicity, the above method suffers from being only first-order in time, regardless of the accuracy of the solvers. Better results can be obtained adopting a different strategy, similar to the Strang splitting described before for the multidimensional case: second-order accuracy is achieved, assuming each subproblem is solved with a method of at least the same accuracy.

Advantages of splitting methods are clear since numerical schemes for the two problems alone are well developed and can be chosen to optimal effect. Despite their advantages, splitting schemes need to be implemented with caution, especially in the choice of operators. There are some situations where this technique does lead to spurious results, or even to wrong numerical solutions: fractional step methods are not adequate, for example, in:

- problems with stiff source terms;
- quasi-steady problems.

Stiff source terms that are not treated carefully can lead to serious numerical difficulties. Computations may produce waves that look reasonable at first glance and yet are propagating at nonphysical speeds due to purely numerical artifacts. The difficulty of solving such problems is that spurious numerical solution phenomena, such as incorrect wave speeds, may occur when insufficient spatial and temporal resolutions are used.

Splitting methods perform very poorly also in those situations where u_t is small relative to the other two terms, in particular when steady or quasi-steady solutions are being sought. For such solutions, highly accurate numerical simulations can only be obtained from numerical methods that respect the balance that occurs between the flux gradient and the source term when u_t is small. In many cases, the fractional step method may not even converge, oscillating in time near the correct solution.

Many strategies have been proposed for solution of these problems. The idea of *source term upwinding* lead Bermúdez and Vázquez-Cendón ([**11**]) to formulate the so-called C-property (for Conservation property), which prevents the propagation of parasitic waves in steady and quasisteady flows. Independently, Greenberg and Leroux ([**32**]) coined the term *well balanced* for schemes that preserve steady states at the discrete level. From these seminal papers, Well Balanced schemes have been explored and developed in various scenarios. Another strategy, described by Gascón and Corberán in [**26**], is to write the source term in divergence form so that it can be incorporated into the flux vector of the homogeneous system to be later discretized in an upwind manner. We refer to [**27**] for a comprehensive list of splitting methods, related properties and literature.

CHAPTER 5

Electrical properties of Single Wall Carbon Nanotubes

The aim of our work is to simulate the system response to the application of an electric potential V: we compute the value of the current $J = J(V)$ generated along the tube by the applied bias. To obtain the induction current, we simulate the evolution of electrons distributions, together with those of phonons to take into account quantum mechanical effects.

Let's recall from Chapter 2 our system of equations, governing the time evolution of particles distributions. The dynamics of electrons and phonons is modeled by the following system of 5 equations:

$$\begin{cases} \partial_t f_i + v_i \partial_x f_i - e_0 E v_F \partial_\varepsilon f_i = \mathcal{C}_i, & i = 1, 2 \\ \partial_t g_\eta + \nu_\eta \partial_x g_\eta = \mathcal{D}_\eta, & \eta = 1, 2, 3, \end{cases} \tag{5.1}$$

where $f_i = f_i(t, x, \varepsilon)$ and $g_\eta = g_\eta(t, x, q)$. We have a two dimensional phase space for f_i's and a two dimensional phase space for g_η's; according to notations in Chapter 3, we have $m = 5$ and $d = 3$; anyway, it will be possible to reduce to a bi-dimensional phase-space in practical computations.

Collision operators at the right hand side can have several different formulations, depending on the accuracy of the considered model and, hence, on the different approximations. Some examples of different approximations and a strategy to compute collision terms in a very efficient way will be described in the sections. We can state here a property all collision terms have in any of the investigated models: they are always given by the sum of two terms, one taking into account interactions between particles of the same type and the other for those between particles of different type (e.g. electron-phonon interactions).

1. Collision terms

The very first models regarding the electrics of CNTs, the latter were considered as one-dimensional quantum wires with ballistic electron transport [**6**]. High-field transport measurement showed that the scattering of electrons with phonons destroys the ballistic behavior and caused reduction of the conductivity at high fields. The generation of phonons in the high-field regime was also confirmed by direct experiments as, for example, Raman scattering measurements.

First models regarding the high-field behavior in metallic SWCNTs were obtained at a macroscopic level or by solving semi-classical Boltzmann equations. In the latter case, the dynamics of electrons was treated in a kinetic way while phonons distributions were kept at a fixed lattice temperature. In this case, only the evolution of electrons needs to be considered, while phonon distribution is kept constant in time. Considering for example the model proposed in [**65**], collision terms were given by the sum of two terms, one accounting for electron-electron interaction given by $\mathcal{C}_i^{ee} = (v_F/l_e)(f_j - f_i)$ and the other accounting

for back-scattering with phonons, given by $\mathcal{C}_i^{ep} = (v_F/l_{pb})[(1-f_i)f_j^+ - f_i(1-f_j^-)]$. Last term represents the gaining ($f_j^+ = f_j(\varepsilon + \hbar\omega)$) or loosing ($f_j^- = f_j(\varepsilon - \hbar\omega)$) of a phonon energy quantum $\hbar\omega$. Constant l_{pb} stands for phonons mean free path; no forward-scattering processes are considered.

It was soon clear that it was necessary to consider also phonons evolutions in order to accurately reproduce nanotubes' electrics. Kinetic equations for phonons distributions were introduced, for example, in [**39, 6**].

In [**39**], electron-electron interactions was similar to that in [**65**] but a more complex electron-phonon collision term was considered, also to take into account the two different types of phonons (in this case, $\eta = 1, 2$) used in this model. Also collision terms for phonons are given by the sum of the phonon-phonon interaction, similar to that for electrons in [**65**], and of the phonon-electron scattering.

In [**6**], the authors proposed an even more accurate model, introducing a third phonon mode in order to describe also forward-scattering processes. Since our simulations are based on the model proposed in this work, we will write explicit formulae for right hand sides of (5.1). Electrons collision operators are given by:

$$\mathcal{C}_i = \mathcal{C}_i^{ac} + \sum_{\eta=1}^{3} \mathcal{C}_i^{\eta}, \tag{5.2}$$

where

$$\mathcal{C}_i^{ac} = \frac{v_F}{l_e}(f_j - f_i), \quad j \neq i, \tag{5.3}$$

models interactions among electrons and

$$\begin{aligned} \mathcal{C}_i^{\eta} \;=\; & \gamma_\eta \left\{ g_\eta(q_i^-) f_j^-(1-f_i) + \left[g_\eta(q_i^+)+1\right] f_j^+(1-f_i) \right. \\ & \left. - g_\eta(q_i^+) f_i(1-f_j^+) - \left[g_\eta(q_i^-)+1\right] f_i(1-f_j^+) \right\} \end{aligned} \tag{5.4}$$

model back ($\eta = 1, 2$) and forward ($\eta = 3$) scattering with phonons. In the above formula, γ_η denotes the electron-phonon coupling constants. We used the abbreviations $f_i^\pm = f_i(t, x, \varepsilon^\pm)$, where

$$\varepsilon^\pm = \varepsilon \pm \hbar\omega_\eta$$

so that $f_i^\pm$ model the emission or absorption of a phonon energy quantum $\hbar\omega_\eta$. The modes $\eta = 1, 2, 3$ refer respectively to K-phonons, longitudinal optical Γ-phonons and transverse optical Γ-phonons. Further, for $\eta = 1, 2$,

$$q_i^\pm = \mp \frac{1}{\hbar v_i} \left(2\varepsilon \pm \hbar\omega_\eta\right), \tag{5.5}$$

while for $\eta = 3$, $q_i^+ = q_i^- = q_i = \omega_3/v_i$. The constant l_{ac} stands for the acoustic mean free path (MFP).

Regarding collision operators for phonons, we have:

$$\mathcal{D}_\eta = \mathcal{D}_\eta^{pp} + \mathcal{D}_\eta^{ep}. \tag{5.6}$$

The first term takes into account phonon-phonon interactions, modeled by

$$\mathcal{D}_\eta^{pp} = -\frac{1}{\tau_\eta}\Big(g_\eta(t, x, q) - g_\eta^0\Big), \tag{5.7}$$

where τ_η denotes the relaxation time and g^0_η is the Bose-Einstein distribution

$$g^0_\eta = \frac{1}{e^{(\hbar\omega_\eta/(k_B\,T))} - 1} \tag{5.8}$$

at a fixed temperature T; $k_B = 8.617\ eV/K$ is the Boltzmann constant. The second term takes into account electron-phonon interactions: for $\eta = 1, 2$ (back scattering)

$$\mathcal{D}^{ep}_\eta = 2\sum_{i=1}^{2} \gamma_\eta \Big\{ (g_\eta + 1) f_i(\varepsilon_i^+)[1 - f_j(\varepsilon_i^-)] - g_\eta\, f_j(\varepsilon_i^-)[1 - f_i(\varepsilon_i^+)] \Big\}, \quad j \neq i, \tag{5.9}$$

where, $f_i(\varepsilon_i^\pm) = f_i(t, x, \varepsilon_i^\pm)$, with

$$\varepsilon_i^\pm = \frac{\hbar}{2}\left(v_i q \pm \omega_\eta\right); \tag{5.10}$$

for $\eta = 3$, instead, electron-phonon collision operator reads

$$\mathcal{D}^{ep}_3 = \gamma_3 \sum_{i=1}^{2} J_i\, \delta_{q,q_i} \int_{\mathbb{R}} \Big\{ (g_3 + 1) f_i(\varepsilon)\left[1 - f_i(\varepsilon^-)\right] - g_3\, f_i(\varepsilon^-)\left[1 - f_i(\varepsilon)\right] \Big\}\, d\varepsilon\,, \tag{5.11}$$

with $\varepsilon^- = \varepsilon - \hbar\omega_3$ and $J_i = 4L/hv_F$ denoting the density of states for electrons of type i with respect to the tube length L. The Kronecker delta δ_{q,q_i} in the collision operator reflects the fact that only phonons with the wave vectors $q_i = \omega_3/v_i$ are emitted and absorbed by forward scattering of electrons.

It should be noted that in all the presented models, scattering lengths are taken as fixed constants, either obtained as fitting parameters or by empiric rules.

2. Constant scattering length

Until phonons are assumed thermalized at room temperature, which means $g_\eta \approx 0$, the contribution of forward scattering has to be neglected (as was the case, e.g., for the model considered in [**65**]). Thus, to compare with experiments, one could only consider backscattering and obtain a simple scaling between scattering length and diameter:

$$l = 65\ d_h.$$

Already in [**40**], authors have pointed out that the assumption of thermalized phonon does not hold: only a significant phonon occupation can, indeed, explain the small value of the measured scattering length l. With a high phonon occupation, both phonon emission and absorption processes are equally relevant, so to take into account more complex scattering processes a better approximation should be considered.

In [**39**], authors considered two different back scattering processes with different scattering lengths: the model was more accurate than the previous one but the considered lengths were too small and the corresponding scattering rate was also too small. The resulting computations were only slightly affected by the scattering lengths and this should not have been the case.

A better approximation was used in [**6**]; their model took into account the role of the time evolution of one more phonon mode, related to forward scattering, and also considered

different values for scattering lengths depending on phonon modes. The quantities l_η, for $\eta = 1, 2, 3$, determining electron-phonon coupling coefficients for the different phonon modes were taken as constants but had, at least, different values: for any phonon mode $\eta = 1, 2, 3$, they depended on tube diameter via:

$$l_1 = 92.0\ d_h \text{ and } l_2 = l_3 = 225.6\ d_h.$$

2.1. Scattering lengths depending on phonon distributions. Looking towards a more realistic model, which could give results comparable with experimental data, we decided it was necessary to consider a variable representation of the scattering lengths, depending on the spatial distribution functions of phonons. What we did, then, was to change from constant values for the EPCs to *variable* ones, computing scattering lenghts in a self-consistent way depending on phonons distributions.

We followed the idea in [**40**] for the computation of the scattering lengths: from now on, we will write the scattering lengths as l, choosing any time the appropriate subscripts depending on the variables used to compute it. A general formula for l is:

$$l_{q\eta} = \frac{\alpha_{q\eta} d_h}{(g_{-q\eta} + 1)}, \tag{5.12}$$

where $g_{q\eta}$ is the phonon occupation. Both g and l depend on q and η; an explicit formula for α in given in [**40**].

In [**40**], authors assumed they could estimate scattering lengths by assuming that the phonon occupation is independent of q and η and by summing the absorption and emission contributions. Scattering lengths are then obtained by substituting a fixed phonon occupation $\bar{g}_0$ in (5.12):

$$l = \frac{65\, d_h}{(1 + \bar{g}_0)}, \tag{5.13}$$

with $\bar{g}_0$ in the $2.7 \sim 5$ range to reconcile the scattering lengths derived from the computed and the measured EPCs. Results obtained with these parameters are consistent with the observation that high-bias saturation currents in SWCNTs on a substrate are significantly higher than those in suspended SWCNTs. Indeed, the effective temperature of optical phonons in suspended SWCNTs is expected to be higher due to the absence of a thermally conductive substrate for heat sinking and such behavior is very well described by this model.

We considered a relation similar to (5.12) but with varying g; first of all we decided to compute EPCs according to the mean in the space variable of phonons distributions: $\bar{g} = \langle g_{q\eta} \rangle_x$. As in [**6**], we assumed similar values for modes $\eta = 2$ and $\eta = 3$.

Although computed current values were comparable with other models, we obtained really different shapes for phonons and electrons distribution functions. For example, our computed distribution for phonon mode $\eta = 1$ was much smaller then in related references.

What we observed was a low transverse optical phonons (the one given by g_3) population having, anyway, very high peaks near the boundaries (see Figure 2.1).

This led us to search for another improvement, considering still different values of scattering lengths for longitudinal optical ($\eta = 2$) and transversal optical ($\eta = 3$) phonon modes. This is theoretically justified by the fact that distribution g_3 takes very small values almost everywhere inside the tube but has very high picks at $q_i^3 = \hbar\omega_3/\hbar v_F$; it makes sense, thus, to

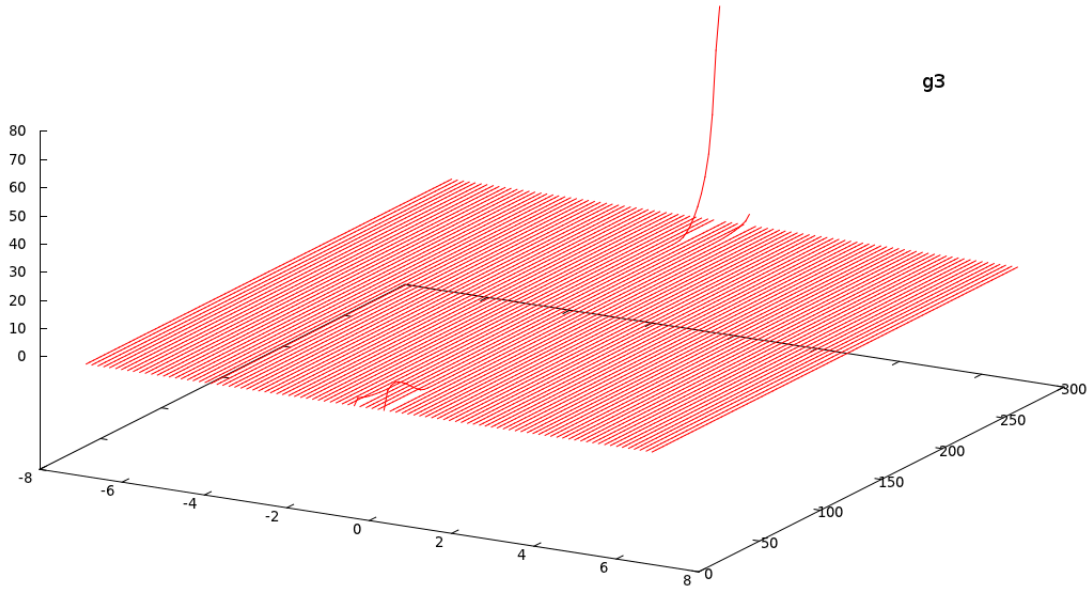

FIGURE 1. Phonon mode $\eta = 3$ distribution.

consider longer paths for longitudinal phonons with respect to those for transverse phonons. We modeled varying EPCs by:

$$l_1 = \frac{92.0\ d_h}{(1+\bar{g}_1)},\ l_2 = \frac{225.6\ d_h}{(1+\bar{g}_2)},\ l_3 = \frac{225.6\ d_h}{(1+\bar{g}_3)}.$$

Simulation results are shown in Figure 2, compared with laboratory results. Thanks to our model, we found great agreement between the computed results for the $J-V$ curves and experimental data ([**16**]). We also observed a significant decrease in computational time needed for the simulations.

2.2. Numerical results for large diameter tubes. We also tried to give a numerical response to a different phenomena reported in [**4**]; this is a comprehensive reference regarding the physics of nanotubes as building blocks in electronic devices and gives important suggestions to create always more realistic models. We applied the presented model to larger diameter nanotubes, which are often used in practical applications.

We used both previous models: the one with constant values for scattering lengths and the one with varying values for scattering lengths for all three kinds of phonons. What we found was that using the former model, current was much less effected by phonons distributions then in the small diameter case; the result was more similar to that of an only ballistic process.

In [**4**] they suggest that *differential conductance*, which is

$$\frac{dV}{dI},$$

should be increasing for large diameter nanotubes. This behavior is opposite with respect to that for small diameter tubes, for which differential conductance decreases for increasing applied voltage. The $I-V$ curves were obtained using the same model simply changing the

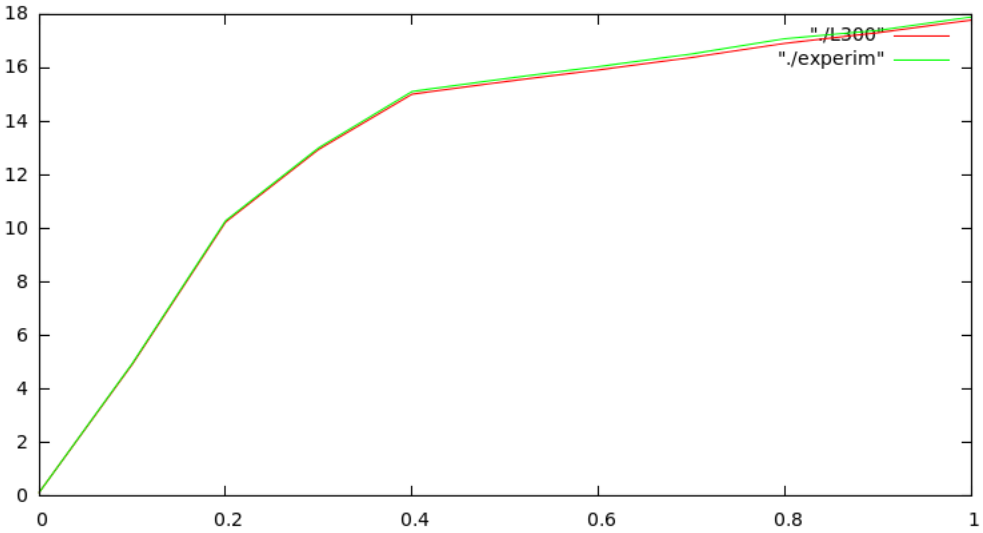

FIGURE 2. $J - V$ curves for computed and experimental data

values of the diameter; the computed results, see Figure 3, show qualitative agreement with the prediction in [**4**] for a wide range of diameters d_h ($6\,nm < d_h < 16\,nm$).

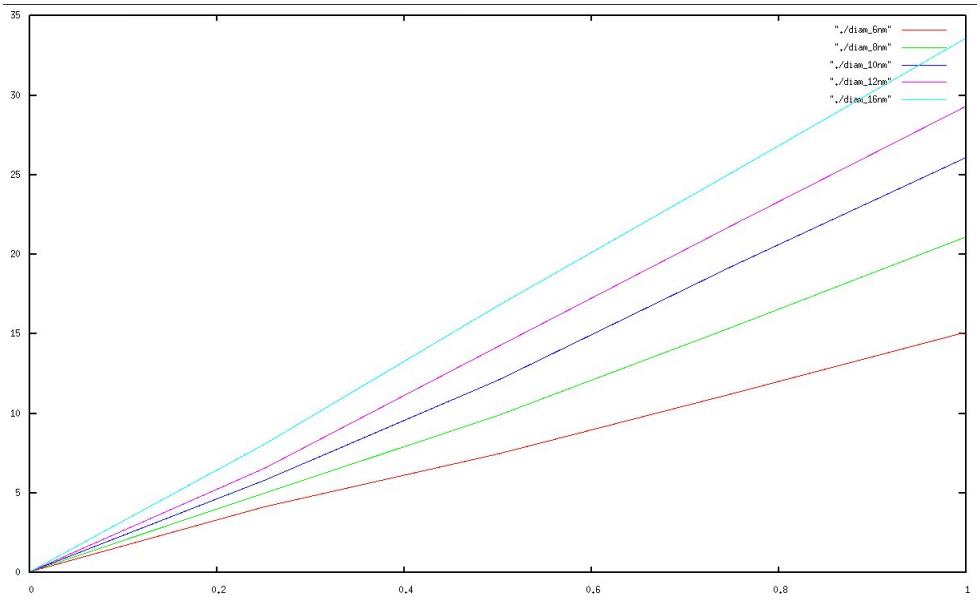

FIGURE 3. $J - V$ curves for large diameter nanotubes ($6\,nm < d_h < 16\,nm$); constant scattering lengths model

Anyway, values of the current are much smaller than the predicted ones ($\sim$ 23% less) meaning this model is not accurate enough for such type of physical problems.

Results obtained using the model with varying scattering lengths gave even worse results (Figure 4).

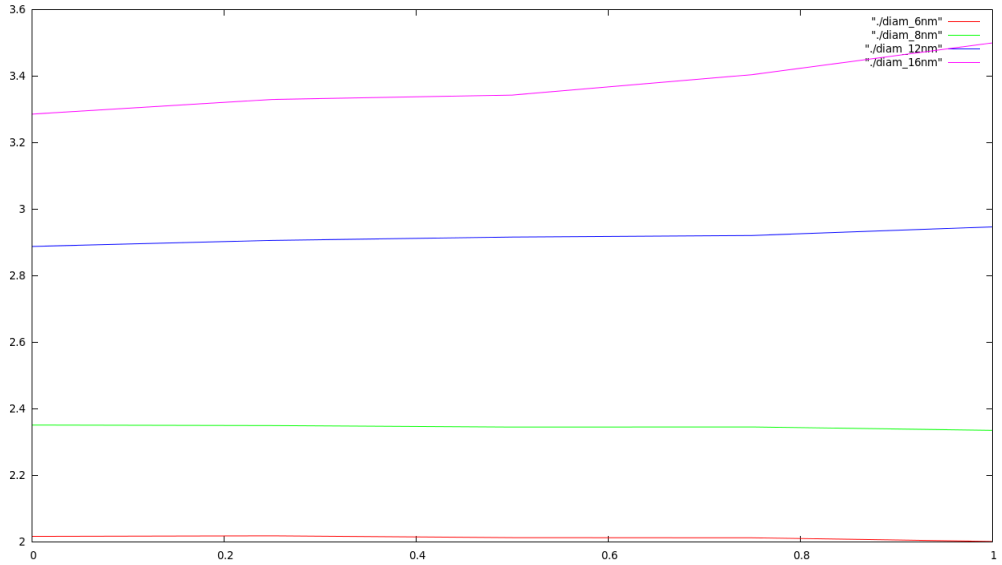

FIGURE 4. $J - V$ curves for large diameter nanotubes ($6\,nm < d_h < 16\,nm$); variable scattering lengths model

Values of the current are almost constant for any applied bias, which is not all a reliable result.

Given the very good results when diameter values are in the correct range, this means a more accurate physical interpretation is needed for larger diameters nanotubes. This will be a very interesting subject for future investigations.

3. Numerical setting

To compute the time evolution of the system we chose a method of line scheme. We first computed approximations to the derivatives of the phase space variables, using a high order (i.e. fifth order) WENO reconstruction in both variables; we approximated the two variables separately by two mono dimensional reconstructions using the global flux splitting technique. For time integration, we used the optimal TVD-RK scheme third order version presented in Chapter 4.

For the numerical approximation of our system we use a fixed uniform discretization for the phase-space variables x, ε and q: for a $L\,nm$ long SWCNT, we define $\Delta x = L/N_x$, where N_x is the chosen number of grid points. In our computations, we considered variable tube lengths: $L = 150\,nm$, $L = 300\,nm$ and $L = 600\,nm$; the best results were obtained for $L = 300\,nm$. Regarding the diameter d_h, we chose the value $d_h = 2$ which is inside the theorical range of validity, i.e. $1\,nm < d_h < 3\,nm$.

For the ε variable it was necessary to make a different choice: the discretization length $\Delta\varepsilon$ of the energy density variable is chosen so that phonon energies $\hbar\omega_\eta$, for $\eta = 1, 2, 3$, are always integer multiples of $\Delta\varepsilon$ for $\eta = 1, 2, 3$, i.e:

$$\hbar\omega_\eta = \sigma_\eta \Delta\varepsilon,$$

with $\sigma_\eta \in \mathbb{N}$. The energy grid is then determined by the values $\varepsilon_n = -\widehat{\varepsilon} + n\Delta\varepsilon$ for $n = 0, \ldots, N_\varepsilon$, with the maximal energy

$$\widehat{\varepsilon} = \frac{N_\varepsilon}{2}\Delta\varepsilon.$$

From this starting point, we define the discretization for the wave vector of phonon mode $\eta = 1, 2, 3$ in the following way:

$$\Delta q = \frac{2\Delta\varepsilon}{\hbar v_F},$$

which gives the grid points

$$q_m^\eta = -\widehat{q}^\eta + m\Delta q, \qquad \text{for } m = 0, \ldots, N_q,$$

where $N_q = N_\varepsilon - \sigma_\eta$ and

$$\widehat{q}^\eta = \frac{N_q}{2}\Delta q = \frac{N_\varepsilon - \sigma_\eta}{2}\Delta q.$$

These discretizations of ε and q guarantee that the following, fundamental *energy and momentum relations*

$$\varepsilon(k') = \varepsilon(k) \pm \hbar\omega_\eta \quad \text{and} \quad k' = k \pm q$$

are satisfied at the discrete level in each individual back-scattering process. Hence, collision operators $\mathcal{C}_i$ and $\mathcal{D}_\eta$ can be evaluated *exactly* in terms of the discretized distribution functions $f_i(t, x, \varepsilon_n)$ and $g_\eta(t, x, q_m^\eta)$. This is a major advantage since no approximations (for example, of extrapolation type) are needed to evaluate collision operators.

We list here explicit formulae for $\varepsilon_i^\pm$ and $q_i^\pm$ which are very useful for all computations. For $\varepsilon_i^\pm$ in 5.10 we have:

$$\begin{aligned}
\varepsilon_1^+ &= \frac{\hbar v_F}{2} q_m + \frac{\hbar\omega_\eta}{2} = \frac{\hbar v_F}{2}\left(-\widehat{q} + m\Delta q\right) + \frac{\sigma_\eta\Delta\varepsilon}{2} = \\
&= -\frac{\hbar v_F}{2}\Delta q \frac{N_\varepsilon - \sigma_\eta}{2} + m\frac{\hbar v_F}{2}\Delta q + \frac{\sigma_\eta\Delta\varepsilon}{2} = \\
&= -\Delta\varepsilon\frac{N_\varepsilon}{2} + \frac{\sigma_\eta\Delta\varepsilon}{2} + m\Delta\varepsilon + \frac{\sigma_\eta\Delta\varepsilon}{2} = \\
&= -\widehat{\varepsilon} + m\Delta\varepsilon + \sigma_\eta\Delta\varepsilon = \varepsilon_m + \sigma_\eta\Delta\varepsilon; \\
\varepsilon_1^- &= \frac{\hbar v_F}{2} q_m - \frac{\hbar\omega_\eta}{2} = \ldots = -\Delta\varepsilon\frac{N_\varepsilon}{2} + \frac{\sigma_\eta\Delta\varepsilon}{2} + m\Delta\varepsilon - \frac{\sigma_\eta\Delta\varepsilon}{2} = \\
&= -\widehat{\varepsilon} + m\Delta\varepsilon = \varepsilon_m;
\end{aligned}$$

$$\begin{aligned}
\varepsilon_2^+ &= -\frac{\hbar v_F}{2} q_m + \frac{\hbar\omega_\eta}{2} = -\frac{\hbar v_F}{2}\left(-\widehat{q} + m\Delta q\right) + \frac{\sigma_\eta\Delta\varepsilon}{2} = \\
&= \frac{\hbar v_F}{2}\Delta q \frac{N_\varepsilon - \sigma_\eta}{2} - m\frac{\hbar v_F}{2}\Delta q + \frac{\sigma_\eta\Delta\varepsilon}{2} = \\
&= \widehat{\varepsilon} - \frac{\sigma_\eta\Delta\varepsilon}{2} - m\Delta\varepsilon + \frac{\sigma_\eta\Delta\varepsilon}{2} = \widehat{\varepsilon} - m\Delta\varepsilon = \varepsilon_m - \Delta\varepsilon; \\
\varepsilon_2^- &= -\frac{\hbar v_F}{2} q_m + \frac{\hbar\omega_\eta}{2} = \ldots = \widehat{\varepsilon} - \frac{\sigma_\eta\Delta\varepsilon}{2} - m\Delta\varepsilon - \frac{\sigma_\eta\Delta\varepsilon}{2} = \\
&= \widehat{\varepsilon} - m\Delta\varepsilon - \sigma_\eta\Delta\varepsilon = \varepsilon_m - \sigma_\eta\Delta\varepsilon.
\end{aligned}$$

For $q_\eta^\pm$ in 5.5 we have:

$$\begin{aligned}
q_1^- &= \frac{1}{\hbar v_F}(2\varepsilon_n - \hbar\omega_\eta) = \frac{2\varepsilon}{\hbar v_F} - \frac{\hbar\omega_\eta}{\hbar v_F} = \frac{-2\widehat{\varepsilon} + 2n\Delta\varepsilon}{\hbar v_F} - \frac{\sigma_\eta \Delta\varepsilon}{\hbar v_F} = \\
&= \frac{-2\Delta\varepsilon}{\hbar v_F}\frac{N_\varepsilon}{2} + n\frac{2\Delta\varepsilon}{\hbar v_F} - \frac{\sigma_\eta}{2}\frac{2\Delta\varepsilon}{\hbar v_F} = -\Delta q\frac{N_\varepsilon}{2} + n\Delta q - \frac{\sigma_\eta}{2}\Delta q = \\
&= -\Delta q\frac{(N_\varepsilon - \sigma_\eta)}{2} + n\Delta q - \frac{\sigma_\eta}{2}\Delta q - \frac{\sigma_\eta}{2}\Delta q = \\
&= -\widehat{q}^\eta + n\Delta q - \sigma_\eta\Delta q = q_n + \sigma_\eta\Delta q; \\
q_1^+ &= -\frac{1}{\hbar v_F}(2\varepsilon_n + \hbar\omega_\eta) = \ldots = \Delta q\frac{(N_\varepsilon - \sigma_\eta)}{2} + \frac{\sigma_\eta}{2}\Delta q + n\Delta q - \frac{\sigma_\eta}{2}\Delta q = \\
&= \widehat{q}^\eta - n\Delta q = -q_n;
\end{aligned}$$

$$\begin{aligned}
q_2^+ &= \frac{1}{\hbar v_F}(2\varepsilon_n + \hbar\omega_\eta) = \frac{-2\widehat{\varepsilon} + 2n\Delta\varepsilon}{\hbar v_F} + \frac{\sigma_\eta\Delta\varepsilon}{\hbar v_F} = -\frac{2\Delta\varepsilon}{\hbar v_F}\frac{N_\varepsilon}{2} + n\Delta q + \frac{\sigma_\eta}{2}\Delta q = \\
&= -\Delta q\frac{(N_\varepsilon - \sigma_\eta)}{2} - \frac{\sigma_\eta}{2}\Delta q + n\Delta q + \frac{\sigma_\eta}{2}\Delta q = -\widehat{q}^\eta + n\Delta q = q_n; \\
q_2^- &= -\frac{1}{\hbar v_F}(2\varepsilon_n - \hbar\omega_\eta) = \ldots = \Delta q\frac{(N_\varepsilon - \sigma_\eta)}{2} + \frac{\sigma_\eta}{2}\Delta q - n\Delta q + \frac{\sigma_\eta}{2}\Delta q = \\
&= \widehat{q}^\eta - n\Delta q + \sigma_\eta\Delta q = -q_n + \sigma_\eta\Delta q.
\end{aligned}$$

Following reference [**7**]. we considered the value $v_F = 8.4\,m/s$ for the Fermi velocity. Regarding phonon energies for K and Γ phonons, we choose the values $\hbar\omega_1 = 160\,meV$ and $\hbar\omega_2 = \hbar\omega_3 = 200\,meV$ respectively, which are good approximations of the real values $\hbar\omega_1 = 161.2\,meV$ and $\hbar\omega_2 = \hbar\omega_3 = 196.0\,meV$, reported in [**6, 7**]. These values for phonon energies $\hbar\omega_\eta$ are integer multiples of $\Delta\varepsilon$, as supposed before, simply choosing $\Delta\varepsilon = 40\,meV$, $\sigma_1 = 4$ and $\sigma_2 = \sigma_3 = 5$.

Regarding group velocities of the optical phonon modes, we chose $\nu_1 = 5000\,m/s$, $\nu_2 = 3000\,m/s$ and $\nu_3 = 0\,m/s$. Velocity ν_1 of the zone-boundary phonons is slightly lower than the value of $7230\,m/s$ given in reference [**7**] at $q = K$. At this symmetry point, however, the group velocity of K phonons takes on its maximum; the lower value $\nu_1 = 5000\,m/s$ used in the transport simulation is a good approximation of the q-averaged phonon velocity.

The electric field E is determined from the applied voltage V by

$$E = \frac{V}{L}.$$

For the relaxation times of the decay of optical phonons due to phonon-phonon interactions, we supposed $\tau_\eta = 3.5\,ps$ for all optical phonon modes $\eta = 1, 2, 3$, which is consistent with experimental values in the literature regarding metallic nanotubes.

Coupling coefficients γ_η for the interaction of electrons with optical phonons depend on scattering lengths l_η ([**6**]) via

$$\gamma_\eta = \frac{v_F}{l_\eta}.$$

Scattering lengths l_η are defined in section 2.2, because they have different formulations depending on the chosen model.

As was already pointed out in Chapter 2, for the electron-electron collision operator 5.3 we considered l_e instead of l_{ac} to take into account electron scattering at impurities.

Regarding boundary conditions we imposed, as it commonly done for hyperbolic PDEs, inflow conditions at the left contact for right moving particles and inflow conditions at the right contact for left traveling particles; this means we assigned

$$\begin{aligned} f_1(t,\,0,\varepsilon) &= t_1^2\, f_0(\varepsilon) + (1-t_1^2) f_2(t,0,\varepsilon) \\ f_2(t,L,\varepsilon) &= t_2^2\, f_0(-\varepsilon) + (1-t_2^2) f_1(t,L,-\varepsilon) \end{aligned}$$

for f_1 and f_2 and

$$g_\eta(t,0,q) = g_\eta^0$$

for g_η's, for phonon modes η such that $\nu_\eta > 0$; g_η^0 is the Bose-Einstein distribution given in (5.8). On respective opposite boundaries, values were simply determined by the time evolution of the system.

The Ohmic contacts were treated as almost perfectly transmitting by assuming $t_i^2 = 0.95$ for the transmission probabilities at the left ($i = 1$) and right ($i = 2$) contacts.

The aim of our simulations was to compute the system response to the application of electric potentials, i.e. the value of the generated current $I(V)$.

We defined the (mean) current at time t as:

$$J(t) = \frac{1}{N_x} \sum_{i=1}^{N_x} J(t, x_i), \tag{5.14}$$

where

$$J(t,x) = \frac{4e_0}{h} \int_{\mathbb{R}} f_2(t,x,\varepsilon) - f_1(t,x,\varepsilon)\, d\varepsilon\,. \tag{5.15}$$

We used the resulting computed current as stopping criteria for our simulations:

$$|J(t+\Delta t) - J(t)| < \varepsilon_J\, |J(t)|.$$

Computations ended when the previous relation was satisfied for $\varepsilon_J = 10^{-3}$.

Regarding ghost grid points, needed for WENO reconstruction, we used extrapolation for ghost terms in the spatial direction while we assigned constant zero value to ghost terms in the energy variable (i.e. for $|\varepsilon| > \widehat{\varepsilon}$).

4. Numerical results for Balance Laws

In the review regarding numerical methods for BLs we presented in Chapter 3, we already pointed out that criteria that ensure reliability of the results for computations regarding difficult BLs systems do not, in general, exist. In our thesis work we tried to find general *validity* conditions for numerical schemes for systems of BLs, at least those similar to our problem.

Computations for the investigated model for CNTs electrics gave physically correct solutions and physically relevant solutions; starting from this successful simulations, we used our numerical scheme for more general problems and tried to find the cases for which it works well.

By extensive computations we found the complete *validity conditions* for the adopted numerical scheme; this means, for example, the range of values of all the parameters for which the scheme computes non oscillatory and (almost) not smeared distributions (which are our PDE solutions).

The numerical methods we choose are very well suited for solving these kind of problems regarding nanotubes because there is not, in such calculations (at least for the present model), the necessity of any change of scale. In this case we are also safe from the very difficult ground of Balance Laws having *stiff* source terms at the right hand side; in such case, methods such as that presented in [**48**] could come in hand. We, thus, prevented our new computations to fall into any of these two categories.

For our study, we worked with adimensional equations and the parameters were left free to assume also non physical values (values not having physical meaning); to keep, however, our study useful to generalize the original model, we considered parameters' variation intervals close to the "real" values. Adimensionalized equations were derived thanks to a change of variables, switching from physical to non physical ones.

Let the new variables be:

$$\tau = \frac{v_F}{L}t, \quad \xi = \frac{x}{L}, \quad \zeta = \frac{\varepsilon}{\widehat{\varepsilon}}, \quad \phi = \frac{q}{\widehat{q}}, \tag{5.16}$$

where L, $\widehat{\varepsilon}$ and $\widehat{q}^{\eta}$ were defined in Section 3. The new distribution functions become:

$$\widehat{f_i}(\tau, \xi, \zeta) = f_i(t(\tau), x(\xi), \varepsilon(\zeta))$$

and

$$\widehat{g}_\eta(\tau, \xi, \phi) = g_\eta(t(\tau), x(\xi), q(\phi)).$$

With some easy calculations, we obtain

$$\frac{v_F}{L}\partial_\tau \widehat{f_i} + \frac{v_F}{L}\partial_\xi \widehat{f_i} - \frac{e_0 E v_F}{\widehat{\varepsilon}}\partial_\zeta \widehat{f_i} = \widehat{\mathcal{C}_i}$$

where $\widehat{\mathcal{C}_i}$ is the collision operator computed in $\widehat{f_i}$ and $\widehat{g}_\eta$. The resulting formulae are:

$$\partial_\tau \widehat{f_i} + \partial_\xi \widehat{f_i} - \widetilde{e}\,\partial_\zeta \widehat{f_i} = \frac{L}{v_F}\widehat{\mathcal{C}_i} \tag{5.17}$$

with

$$\widetilde{e} = \frac{e_0 E L}{\widehat{\varepsilon}}. \tag{5.18}$$

Similarly, $\widehat{g}_\eta$ become, for $\eta = 1, 2, 3$,

$$\frac{v_F}{L}\partial_\tau \widehat{g}_\eta + \frac{\nu_\eta}{L}\partial_\xi \widehat{g}_\eta = -\frac{1}{\tau_\eta}(\widehat{g}_\eta - g_\eta^0) + \widehat{\mathcal{D}}_\eta^{ep},$$

from which we obtain

$$\partial_\tau \widehat{g}_\eta + \widetilde{\nu}_\eta \partial_\xi \widehat{g}_\eta = -\frac{L}{v_F \tau_\eta}(\widehat{g}_\eta - g_\eta^0) + \frac{L}{v_F}\widehat{\mathcal{D}}_\eta^{ep}, \tag{5.19}$$

where

$$\widetilde{\nu}_\eta = (\nu_\eta / v_F). \tag{5.20}$$

We let all parameters (e.g. $\widetilde{e}$, $\widetilde{\nu}_\eta$) vary in intervals centered near the "reference values", i.e. the values of the parameters in (5.18) and (5.20). As an example, we have $\widetilde{e} \approx O(1)$ while $\widetilde{\nu}_\eta \approx O(10^{-2})$ for $\eta = 1, 2$; similar relations hold for scattering times at the right hand sides. We also decided to hold the CFL bound at a fixed value, the one that it has in the initial problem with equations (5.17) and (5.19), thus not depending on the varying parameters; such choice caused all *validity intervals* being not centered around the starting value.

Since we had no meaningful physical assumptions to use as stopping criteria, a reasonable assumption that could tell us when to stop calculations, we decided to compute *stationary* solutions. It was necessary, then, to consider a different type of WENO solver, one that could be appropriate in this case. For this porpoise, we considered a high order *Well-Balanced* scheme, based again on a finite difference WENO reconstruction ([**64**]); we did not change the time solver. We decided to discard solutions, i.e. the corresponding set of parameters, when (big) oscillations or (too much) smearing started to appear in the computed solutions or if convergence was not obtained before 500 time steps.

We found the considered numerical strategy gives good results, i.e. (essentially) non oscillatory and little smeared solutions, for a quite large set of parameters. We wish to give a general theorical explanation for the obtained results, which could explain the validity of the considered scheme and, thus, be a general and reliable reference scheme for problems related to, quite general, hyperbolic Balance Laws.

Bibliography

[1] http://nobelprize.org/nobel_prizes/physics/laureates/2010/press.html.

[2] http://www.zyvex.com/nanotech/feynman.html, 1959.

[3] N. Allinger. Conformational analysis. 130. mm2. a hydrocarbon force field utilizing v1 and v2 torsional terms. *J. Am. Chem. Soc.*, 99:8127–8134, 1977.

[4] M.P. Anantram and F. Léonard. Physics of carbon nanotube electronic devices. *Rep. Prog. Phys.*, 69:507–561, 2006.

[5] M. Arroyo and T. Belytschko. Finite element methods for the non-linear mechanics of crystalline sheets and nanotubes. *Int. J. Num. Meth. Eng.*, 59:419–456, 2004.

[6] C. Auer, F. Schürrer, and C. Ertler. Deterministic solution of boltzmann equations governing the dynamics of electrons and phonons in cabon nanotubes. In L. Puccio V. Cutello, G. Fotia, editor, *Applied and Industrial Mathematics in Italy, II*, pages 89–100. World Scientific, Singapore, 2006.

[7] C. Auer, F. Schürrer, and C. Ertler. Hot phonon effects on the high-field transport in metallic carbon nanotubes. *Phys. Rev. B*, 74:165409–165419, 2006.

[8] C. Auer, F. Schürrer, and C. Ertler. Influence of hot phonons on the transport properties of single wall carbon nanotubes. *J. Comput. Electron.*, 6:325–328, 2007.

[9] P. Avouris. Carbon nanotube electronics. *Chemical Physics*, 281:429–445, 2002.

[10] R. Bawa, S.R. Bawa, S.B. Maebius, T.Flynn, and C. Wei. Protecting new ideas and inventions in nanomedicine with patents. *Nanomedicine: Nanotechnology, Biology and Medicine*, 1:150–158, 2005.

[11] A. Bermúdez and M.E. Vázquez. Upwind methods for hyperbolic conservation laws with source terms. *Computers and Fluids*, 23:1049–1071, 1994.

[12] H.P. Boehm, A. Clauss, G.O. Fischer, and U. Hofmann. Das adsorptionsverhalten sehr dünner kohlenstoffolien. *Zeitschrift für anorganische und allgemeine Chemie*, 316:119–127, 1962.

[13] D. Brenner, O. Shenderova, J. Harrison, and S. Sinnott B. Ni. A second-generation reactive empirical bond order (rebo) potential energy expression for hydrocarbons. *J. Phys: Condensed Matter*, 14:783–802, 2002.

[14] G. Busetto and M. Morandi Cecchi. Computational mechanical modeling of the behavior of carbon nanotubes. In *Proceedings of the 7th Conference on Systems Theory and Sc. Comp.*, pages 58–65. World Scientific and Engineering Academy and Society, 2007.

[15] D. Caillerie, A. Mourad, and A. Raoult. Discrete homogenization in graphene sheet modeling. *J. Elast.*, 84:33–68, 2006.

[16] M. Morandi Cecchi and V. Rispoli. Numerical solution of electrons' and phonons' coupled dynamics in carbon nanotubes. *Communications in Applied and Industrial Mathematics*, Submitted, 2011.

[17] M. Morandi Cecchi, V. Rispoli, and M. Venturin. An atomic-scale finite element method for single wall carbon nanotubes. In V. Valente E. De Bernardis, R. Spigler, editor, *Applied and Industrial Mathematics in Italy, III*, pages 449–460. World Scientific, Singapore, 2006.

[18] T. Chang and H. Gao. Size-dependent elastic properties of a single walled carbon nanotube via a molecular mechanics model. *J. Mech. and Phys. Solids*, 51:1059–1074, 2003.

[19] J.C. Charlier, X. Blase, and S. Roche. Electronic and transport properties of nanotubes. *Rev. Mod. Phys.*, 79:677–732, 2007.

[20] R. Courant, K.0. Friedrichs, and H. Lewy. Uber die partiellen differenzengleichungen der mathematisches physik. *Math. Ann.*, 100:32–74, 1928.

[21] M. G. Crandall and A. Majda. The method of fractional steps for conservation laws. *Math. Comp.*, 34:285–314, 1980.

[22] C.M. Dafermos. *Hyperbolic Conservation Laws in Continuum Physics*. Springer, 1999.
[23] M.S. Dresselhaus, G. Dresselhaus, and A. Jorio. *Group Theory - Application to the Physics of Condensed Matter*. Springer-Verlag Berlin, 2008.
[24] E.B. Barrosa et al. Review on the symmetry-related properties of carbon nanotubes. *Physics Reports*, 431:261–302, 2006.
[25] K.S. Novoselov et al. Electric field effect in atomically thin carbon films. *Science*, 306:666–669, 2004.
[26] Ll. Gascón and J.M. Corberán. Construction of second-order tvd schemes for nonhomogeneous hyperbolic conservation laws. *J. Comput. Phys.*, 172:261–297, 2001.
[27] A.M. Gavarra. *High Resolution Schemes for Hyperbolic Conservation Laws with Source Terms*. PhD thesis, University of Valencia, 2009.
[28] E. Godlewski and P. A. Raviart. *Numerical Approximation of Hyperbolic System of Conservation Laws*. Springer, 1996.
[29] S.K. Godunov. *Mat. Sb.*, 47:271, 1959.
[30] J. B. Goodman and R. J. LeVeque. On the accuracy of stable schemes for 2d scalar conservation laws. *Math. Comp.*, 45:15–21, 1985.
[31] S. Gottlieb and C.W. Shu. Total variation diminishing runge-kutta schemes. *Mathematics of Computation*, 67:73–85, 1998.
[32] J.M. Greenberg and A.Y. Leroux. A well-balanced scheme for the numerical processing of source terms in hyperbolic equations. *SIAM J. of Num. Anal.*, 33:1–16, 1996.
[33] A. Harten, B. Engquist, S. Osher, and S. Chakravarthy. Uniformly high order essentially non-oscillatory schemes, iii. *J. of Computational Physics*, 71:231–303, 1987.
[34] S. Iijima. Helical microtubules of graphitic carbon. *Nature*, 354:56–58, 1991.
[35] G. Jiang and C.W. Shu. Efficient implementation of weighted eno schemes. *J. of Computational Physics*, 126:202–228, 1996.
[36] K.N. Kudin, G.E. Scuseria, and B.I. Yakobson. C_2f, bn, and c nanoshell elasticity from ab-initio computations. *Phys. Rev. B*, 64:235406–235415, 2001.
[37] J. Kürti, G. Kresse, and H. Kuzmany. First-principles calculations of the radial breathing mode of single wall carbon nanotubes. *Phys. Rev. B*, 58:8869–8872, 1998.
[38] P.D. Lax and B. Wendroff. Systems of conservation laws. *Comm. Pure Appl. Math.*, 13:217–237, 1960.
[39] M. Lazzeri and F. Mauri. Coupled dynamics of electrons and phonons in metallic nanotubes: current saturation from hot phonon generation. *Phys. Rev. B*, 73:165419–165424, 2006.
[40] M. Lazzeri, S. Piscanec, F. Mauri, A.C. Ferrari, and J. Robertson. Electron transport and hot phonons in carbon nanotubes. *Phys. Rev. Lett.*, 95:236802–236805, 2005.
[41] R. J. LeVeque. *Time-split methods for partial differential equations*. PhD thesis, Stanford, 1982.
[42] R.J. LeVeque. *Numerical Methods for Conservation Laws*. Lectures in Mathematics. Birkhäuser Verlag, 1992.
[43] B. Liu, Y. Huang, H. Jiang, and K. Hwang S. Qu. The atomic-scale finite element method. *Comp. Meth. Appl. Mech. Eng.*, 193:1849–1864, 2004.
[44] X.D. Liu, S. Osher, and T. Chan. Weighted essentially nonoscillatory schemes. *J. of Computational Physics*, 115:200–212, 1994.
[45] J. Lu. Elastic properties of carbon nanotubes and nanoropes. *Phys. Rev. Lett.*, 79:1297–1300, 1997.
[46] A. Oberlin, M. Endo, and T. Koyama. Filamentous growth of carbon through benzene decomposition. *J. Crystal Growth*, 32:335–349, 1976.
[47] S. Osher. Riemann solvers, the entropy condition, and difference approximations. *SIAM J. Num. Anal.*, 21:217–235, 1984.
[48] L. Pareschi and G. Russo. Implicit-explicit runge-kutta methods and applications to hyperbolic systems with relaxation. *J. Sci. Comput.*, 25:129–155, 2005.
[49] E. Pop, D. Mann, J. Cao, Q. Wang, K. Goodson, and H. Dai. Negative differential conductance and hot phonons in suspended nanotube molecular wires. *Phys. Rev. Lett.*, 95:155505, 2005.
[50] D. Qian, G. Wagner, and W. Liu. A multiscale projection method for the analysis of carbon nanotubes. *Comp. Meth. Appl. Mech. Eng.*, 193:1603–1632, 2004.

[51] L.V. Radushkevich and V.M. Lukyanovich. About the structure of carbon formed by thermal decomposition of carbon monoxide on iron substrate. *Zurn. Fisic Chim.*, 26:88–95, 1952.

[52] R.D. Richtmyer and K.W. Morton. *Difference Methods for Initial-value Problems*. Wiley-Interscience, 1967.

[53] V. Rispoli. Carbon nanotubes modeling by an algebraic point of view. *GNCS Congress*, Presented speech, 2009.

[54] P.L. Roe. Approximate riemann solvers, parameter vectors, and difference schemes. *J. of Computational Physics*, 43:357–372, 1981.

[55] R. Saito, M.S. Dresselhaus, and G. Dresselhaus. *Physical properties of carbon nanotubes*. Imperial College Press, London, 1998.

[56] D. Sanchez-Portal, E. Artacho, J. Soler, A. Rubio, and P. Ordejon. Ab-initio structural, elastic and vibrational properties of carbon nanotubes. *Phy. Rev. B*, 59:12678–12688, 1999.

[57] C.W. Shu. Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws. *NASA/CR-97-206253*, ICASE Report No. 97-65:1–78, 1997.

[58] C.W. Shu and S. Osher. Efficient implementation of essentially non-oscillatory shock capturing schemes. *J. of Computational Physics*, 77:439–471, 1988.

[59] C.W. Shu and S. Osher. Efficient implementation of essentially non-oscillatory shock capturing schemes ii. *J. of Computational Physics*, 83:32–78, 1989.

[60] G. Strang. Accurate partial difference methods ii: nonlinear problems. *Numer. Math.*, 6:37–46, 1964.

[61] G. Strang. On the construction and comparison of difference schemes. *SIAM J. Num. Anal.*, 5:506–517, 1968.

[62] Z. K. Tang, L. Zhang, N. Wang, X.X. Zhang, G.H. Wen, G.D. Li, J.N. Wang, C.T. Chan, and P. Sheng. Superconductivity in 4 angstrom single-walled carbon nanotubes. *Science*, 292:2462–2465, 2001.

[63] Z.C. Tu and Z.C. Ou-Yang. Single-walled and multi-walled carbon nanotubes viewed as elastic tubes with the effective young's moduli dependent on layer number. *Phys. Rev. B*, 65:233407–233410, 2002.

[64] Y. Xing and C.W. Shu. High-order well-balanced finite difference weno schemes for a class of hyperbolic systems with source terms. *J. Sci. Comput.*, 27:477–494, 2006.

[65] Z. Yao, C. Kane, and C. Dekker. High-field electrical transport in single-wall carbon nanotubes. *Phys. Rev. Lett.*, 84:2941–2944, 2000.

[66] X. Zhou, H. Chen, J. Zhou, and O.Y. Zhong-Can. The structure relaxation of carbon nanotube. *Physica B*, 304:86–90, 2001.