Turbulence & Wind Energy

by ***E. Plaut, J. Peinke & M. Hölling*** at Mines Nancy

Version of December 6, 2022

Contents

Introduction

This is the latest / current version of the lecture notes of Emmanuel Plaut, Joachim Peinke and Michael Hölling for the module ***Turbulence & Wind Energy*** at **Mines Nancy**, in the Department Energy, at the Master 2 Level.

This module is somehow a follow up of the modules ***Mécanique des fluides 2 : ondes, couches limites et turbulence*** (Plaut 2021*a*) and ***Turbomachines - Énergies hydraulique et éolienne*** (Jenny 2017) of the Master 1 Level. Indeed, one can find, in the last chapter of Plaut (2021*a*), an introduction to turbulent flows; in the chapter 4 of Jenny (2017), an introduction to wind energy.
Since, however, not every students following this lecture has followed these previous lectures, we restart from the level of a basic Master 1 Level fluid mechanics lecture. The chapter 0 is there to remind the ***base fluid mechanics equations***, under the ***continuum mechanics approach***, and to fix some notations.
Importantly, we only consider ***incompressible fluids*** in this lecture.

In the first part, that corresponds to the chapter 1, driven by E. P., we study ***turbulence modelling for computational fluid dynamics*** **(CFD)**.
We focus on the ***Reynolds Averaged Navier-Stokes*** **(RANS)** ***approach*** and ***models***. The RANS approach is closely linked to the ***statistical theory of turbulence***, which is quite relevant. The RANS approach is basically interesting to characterize and gain some understanding on ***turbulence phenomena***. Finally, RANS models are still, today, the preferred choice for engineering studies. To place this in the History of science with a capital H, note that the concept of *Reynolds Average* has been introduced in Reynolds (1895), who was, in fact, more concerned with the *transition to turbulence* than with *turbulence* itself[1]. To write the chapter 1 of the present document, we have been inspired by Hanjalić & Launder (2011); Wilcox (2006).
During the corresponding sessions (1 to 3), using a *problem-based learning approach*, studies of the ***Direct Numerical Simulations*** **(DNS)** ***database*** of Lee & Moser (2015) will be performed with **Matlab**.

In the second part, driven by J. P. & M. H., that corresponds to the chapter 2 (and appendix A), we do give some insights on the ***statistical theory of turbulence***, concentrating on the fluctuations (not precisely characterized in Plaut 2021*a*), and on applications to ***wind energy***. We thus also review in this part the ***aerodynamics of wind turbines***.
During the corresponding sessions (4 to 6), studies of various ***experimental turbulence data*** will be performed with the free software **R**, dedicated to statistical computing[2].

[1] More modestly, to place this in the syllabus of our Master 2, note that Plaut (2021*b*) cited Reynolds (1895)...

[2] See `www.r-project.org` .

The aim of this document is to give a framework for the lectures, and problems and exercises that will be solved 'interactively' during the sessions. Some notes in the chapter 2 and appendix A are only devoted to the most interested readers... To save paper, ink and energy, the printed version of this document does not include this appendix; it may be read only in the PDF version of the lecture notes available on the ***dynamic web page of this module***

http://emmanuelplaut.perso.univ-lorraine.fr/twe .

This web page also contains turbulent flow data sets, that have been given by J. P., for exercise 2.3... For the students of Mines Nancy, this module has an Intranet - ARCHE page

http://arche.univ-lorraine.fr/course/view.php?id=11599

on which wind speed and wind turbine power data, grid turbulence data are available for exercises 2.2 and 2.3.

In the chapter 1, many equations have some ***blank spaces***, e.g., equation (1.56). The student should write by himself the solutions here, during, or, after the lectures. Accordingly, there are many ***empty or incomplete figures***, e.g. figure 1.2, signaled by a 'DIY' in the legend, which means 'Do it yourself'. These figures should in principle be realized by the student. A more complete version of this chapter, where there are less blank spaces and empty figures, will be published at the end of session 3, and may be asked by mail to E. P.
The solutions of the exercises and problems solved during the sessions will be displayed on 'video presentations' posted on the web page of the module after the sessions. It is quite important that the students work on these 'video presentations', since they propose figures and physical interpretations that are often not included in these lecture notes.

E. P. thanks Boris Arcen, J. P. & Carlos Peralta for interesting discussions and inputs, that helped him to dive into the topics of turbulence. He also thanks Matthieu Gisselbrecht for the cooperation of year 2015 and Stefan Heinz for the cooperation running since 2017.
J. P. & M. H. acknowledge support by the *Fondation Mines Nancy* and *Erasmus* + program.

Nancy & Oldenburg, December 6, 2022.

Emmanuel Plaut, Joachim Peinke & Michael Hölling.

Chapter 0

The basic concepts and equations of fluid dynamics - Notations

0.1 Cartesian coordinates, tensors and differential operators

We most often use a ***cartesian system of coordinates*** of origin O, associated to the laboratory frame. The coordinates are denoted (x, y, z) or (x_1, x_2, x_3), and we use Einstein's convention of summation over repeated indices, e.g. the position vector is

$$\mathbf{x} \;=\; x_i \mathbf{e}_i$$

with $\mathbf{e}_1,\ \mathbf{e}_2,\ \mathbf{e}_3$ the orthonormal base vectors. We do not use overbars to designate vectors, since the overbars will be used to define various averages, e.g. in chapter 1, 'Reynolds averages'. Vectors and tensors of order 2 will thus be simply denoted by boldface characters. The use of tensorial intrinsic notations will be minimized, to save some energy to face other difficulties. However, it would be interesting to reformulate all the equations written in cartesian coordinates using tensorial intrinsic notations, and this task, as an intermediate step, would be quite important in fact to study e.g. turbulent pipe flow with, at the end, cylindrical coordinates.
Last but not least, ∂_t (resp. ∂_{x_i}) denotes the differential operator that takes the partial derivative with respect to the time t (resp. coordinate x_i).

0.2 Dynamics of incompressible newtonian fluids

The ***eulerian velocity field*** $\mathbf{v}(\mathbf{x}, t)$ is used to describe the flow of ***incompressible newtonian fluids***. The ***incompressibility*** means that the ***mass density*** ρ is uniform and constant. Therefore the ***mass conservation equation*** reads

$$\boxed{\partial_{x_i} v_i \;=\; 0}\ , \tag{0.1}$$

which means that the velocity field $\mathbf{v}$ is conservative.
The ***Cauchy stress tensor*** $\boldsymbol{\sigma}$ determines the surface force $d^2\mathbf{f}$ exerted on a small surface of area d^2A and normal unit vector $\mathbf{n}$ pointing outwards, by the exterior onto the interior, through

$$d^2\mathbf{f} \;=\; \boldsymbol{\sigma} \cdot \mathbf{n}\; d^2A \;=\; \sigma_{ij} n_j\; d^2A\; \mathbf{e}_i\ . \tag{0.2}$$

The stresses σ_{ij} are given by the sum of ***pressure*** and ***viscous stresses***,

$$\sigma_{ij} \;=\; -\,p_{\text{static}}\delta_{ij} \;+\; 2\eta S_{ij}(\mathbf{v}) \tag{0.3}$$

with p_{static} the ***static pressure***, η the ***dynamic viscosity*** of the fluid, $\mathbf{S}(\mathbf{v})$ the ***rate-of-strain tensor*** defined as the symmetric part of the velocity gradient, i.e., in components,

$$\boxed{S_{ij}(\mathbf{v}) \;:=\; \frac{1}{2}(\partial_{x_i} v_j + \partial_{x_j} v_i)} \,. \tag{0.4}$$

In this equation, the sign := means a definition.
The linear momentum equation is the ***Navier-Stokes equation***

$$\rho[\partial_t v_i \;+\; \partial_{x_j}(v_i v_j)] \;=\; \rho g_i \;+\; \partial_{x_j}\sigma_{ij} \;=\; \rho g_i - \partial_{x_i} p_{\text{static}} \;+\; \partial_{x_j}(2\eta S_{ij}(\mathbf{v})) \;, \tag{0.5}$$

where g_i are the components of the acceleration due to gravity. Usually, a ***modified pressure*** that includes a gravity term,

$$p \;=\; p_{\text{static}} \;+\; \rho g Z \;, \tag{0.6}$$

where Z is a vertical coordinate, is used, to group the first two terms on the r.h.s. of equation (0.5), which reads therefore

$$\boxed{\rho[\partial_t v_i \;+\; \partial_{x_j}(v_i v_j)] \;=\; -\,\partial_{x_i} p \;+\; \partial_{x_j}(2\eta S_{ij}(\mathbf{v}))} \,. \tag{0.7}$$

In the chapter 1, we will most often use the modified pressure p instead of the static pressure p_{static} .
After dividing by the mass density, we get another form of the ***Navier-Stokes equation***,

$$\boxed{\partial_t v_i \;+\; v_j \partial_{x_j} v_i \;=\; -\,\frac{1}{\rho}\partial_{x_i} p \;+\; \partial_{x_j}(2\nu S_{ij}(\mathbf{v}))} \tag{0.8}$$

with $\nu = \eta/\rho$ the ***kinematic viscosity*** of the fluid. We have used the mass conservation equation (0.1), $\partial_{x_j} v_j = 0$, to rewrite the nonlinear term on the l.h.s.[1]; this whole l.h.s. is in fact the acceleration of the fluid particle (in the sense of the continuum mechanics) that passes through $\mathbf{x}$ at time t...
For the sake of simplicity, we have only considered, as volume forces, the gravity forces, and have 'hid' them into the modified pressure. It may be interesting to consider other forces; this is done in Hanjalić & Launder (2011)...

[1] An equation reads generically l.h.s. = r.h.s. with l.h.s. its 'left hand side', r.h.s. its 'right hand side'.

Chapter 1

Reynolds Averaged Navier-Stokes approach and models

This chapter corresponds to the sessions 1 *to* 3.

We start with an introduction of the ***Reynolds approach*** and then go quickly to the ***Reynolds equations and stresses***, which are the common bases of all ***Reynolds Averaged Navier-Stokes* (RANS) *models***. We discuss briefly the ***Reynolds stress equations*** and ***models***, and then focus on ***eddy-viscosity 2-equations models***, namely, the $k-\varepsilon$ and $k-\omega$ ***models***. In a *problem-based learning approach*, we propose the problems 1.1 and 1.2 to dive into this subject with analytical calculations and numerical comparisons with a reference database, that offers online results of ***direct numerical simulations* (DNS)** of ***channel flow***. Precisely, we will rummage with **Matlab**[1] in the DNS database of the University of Texas at Austin,

`https://turbulence.oden.utexas.edu` . (1.1)

The DNS on which we will focus are presented in Lee & Moser (2015).

1.1 Reynolds decomposition, equations and stresses

1.1.1 Reynolds average and decomposition

A natural and most important approach in turbulence is the ***statistical approach***. A turbulent flow experiment (possibly, a numerical experiment !) may be repeated N times, with N a large integer, and one may average for instance the pressure field $p(\mathbf{x},t)$ over all the realizations. Denoting $p^n(\mathbf{x},t)$ the pressure field at position $\mathbf{x}$ and time t during the n^{th} realization, the ***ensemble average*** or ***Reynolds average*** of the pressure is

$$\boxed{\overline{p}(\mathbf{x},t) \;=\; \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} p^n(\mathbf{x},t)} \;. \tag{1.2}$$

Often, the turbulence is ***'stationary'***, and one may, with a reasonable hypothesis of ***ergodicity***, use alternately a ***time average*** to define the ***mean pressure***

$$\overline{p}(\mathbf{x}) \;=\; \langle p(\mathbf{x},t)\rangle_t \;=\; \lim_{T\to\infty}\frac{1}{T}\int_0^T p(\mathbf{x},t)\;dt \tag{1.3}$$

[1] Another software may of course be used !

where, on the r.h.s., in the integral, it might be relevant to write p^1 instead of p, since we refer to one realization (say, the first one, labelled $n = 1$) of the flow.
The ***Reynolds decomposition*** splits the pressure p and the velocity fields $\mathbf{v}$ into ***mean values*** + ***fluctuations***, according to

$$\begin{array}{lccccc} & \boxed{p \;=\; P \;+\; p'} & , & \boxed{v_i \;=\; U_i \;+\; u_i} & \Longleftrightarrow & \boxed{\mathbf{v} \;=\; \mathbf{V} \;+\; \mathbf{u}} \\ \text{with} & P \;=\; \overline{p} & , & U_i \;=\; \overline{v_i} & \Longleftrightarrow & \mathbf{V} \;=\; \overline{\mathbf{v}}\,, \\ \text{hence} & \overline{p'} = 0 & , & \overline{u_i} \;=\; 0 & \Longleftrightarrow & \overline{\mathbf{u}} \;=\; \mathbf{0}\,. \end{array} \tag{1.4}$$

1.1.2 Reynolds equations and stresses or second moments

As stated in the introduction, the fluid is assumed to be ***incompressible***, hence the ***mass conservation equation*** is given by (0.1),

$$\partial_{x_i} v_i \;=\; 0\,. \tag{1.5}$$

From the definition (1.2), it is clear the one has the ***commutation rules***

$$\overline{\partial_t p} \;=\; \partial_t \overline{p} \quad \text{and} \quad \forall i\,, \quad \overline{\partial_{x_i} p} \;=\; \partial_{x_i} \overline{p}\,. \tag{1.6}$$

This applies to any real-valued field, i.e. also to the velocity components. Therefore, by taking the Reynolds average of the equation (1.5), we obtain

$$\overline{\partial_{x_i} v_i} \;=\; \partial_{x_i} \overline{v_i} \;=\; 0 \quad \Longleftrightarrow \quad \boxed{\partial_{x_i} U_i \;=\; 0}\,, \tag{1.7}$$

i.e. the mean flow is also conservative. This equation (1.7) is the **1st** ***RANS equation***. Since the mass conservation equation (1.5) is linear, its average form (1.7) is exactly the same.

Let us now consider the ***Navier-Stokes equation*** (0.7),

$$\rho[\partial_t v_i \;+\; \partial_{x_j}(v_i v_j)] \;=\; -\,\partial_{x_i} p \;+\; \partial_{x_j}(2\eta S_{ij}(\mathbf{v}))\,. \tag{1.8}$$

On purpose, the nonlinear advection term on the l.h.s. has been written under a 'conservative' form. When one takes the Reynolds average of the equation (1.8), the averages of the linear terms yield no surprise according to the commutation rules (1.6), i.e. we get

$$\rho\big(\partial_t U_i \;+\; \partial_{x_j} \overline{v_i v_j}\;\big) \;=\; -\,\partial_{x_i} P \;+\; \partial_{x_j}(2\eta S_{ij}(\mathbf{V}))\,. \tag{1.9}$$

On the contrary, the nonlinear term has to be treated carefully. Indeed

$$\overline{v_i v_j} \;=\; \overline{(U_i + u_i)(U_j + u_j)} \;=\; \overline{U_i U_j} \;+\; \overline{U_i u_j} \;+\; \overline{u_i U_j} \;+\; \overline{u_i u_j} \;=\; U_i U_j \;+\; U_i \overline{u_j} \;+\; \overline{u_i} U_j \;+\; \overline{u_i u_j}\,.$$

To transform the second and third terms on the r.h.s., we made use of the fact that $V = U_i$ or U_j do not depend on the realization, hence, with $w = u_j$ or u_i ,

$$\overline{Vw} \;=\; \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} V(\mathbf{x},t)\; w^n(\mathbf{x},t) \;=\; V(\mathbf{x},t) \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} w^n(\mathbf{x},t) \;=\; V(\mathbf{x},t)\; \overline{w}(\mathbf{x},t) \;=\; V\; \overline{w}\,.$$

Since $\mathbf{u}$ is a fluctuating vector field of zero average by its definition (1.4), we get

$$\overline{v_i v_j} \;=\; U_i U_j \;+\; \overline{u_i u_j} \quad \Longleftrightarrow \quad \overline{u_i u_j} \;=\; \overline{v_i v_j} \;-\; \overline{v_i}\;\overline{v_j}\,. \tag{1.10}$$

Using the terminology of statistics, the first of the equations (1.10) reads

$$\text{mean}(v_i v_j) \;=\; (\text{mean}(v_i))\;(\text{mean}(v_j)) \;+\; \text{covariance}(v_i v_j)\;. \tag{1.11}$$

Most importantly, if the fluctuations u_i and u_j are correlated, the ***second moment***

$$\boxed{\overline{u_i u_j} \;=\; \text{covariance}(v_i v_j)} \tag{1.12}$$

does not vanish. Thus we get a 'fluctuation-induced' contribution in the ***RANS equation*** (1.9), that is put on the r.h.s.,

$$\boxed{\rho[\partial_t U_i \;+\; \partial_{x_j}(U_i U_j)] \;=\; -\,\partial_{x_i} P \;+\; \partial_{x_j}(2\eta S_{ij}(\mathbf{V})) \;+\; \partial_{x_j}\tau_{ij}} \tag{1.13}$$

with the ***Reynolds stresses***

$$\boxed{\tau_{ij} \;=\; -\,\rho\overline{u_i u_j} \;=\; -\,\rho\;\text{covariance}(v_i v_j)}\;. \tag{1.14}$$

The equation (1.13) is the $\mathbf{2^{ème}}$ ***RANS equation***, or the ***RANS momentum equation***. It shows that, if the fluctuations u_i and u_j are correlated, they feedback onto the mean-flow through the term τ_{ij} that corrects the viscous stress $2\eta S_{ij}(\mathbf{V})$ (see the equation 0.3) in the evolution equation for U_i ; this 'analogous role' explains the terms 'Reynolds stress' or 'Reynolds stresses', though there are no surface forces (see the equation 0.2) behind this.
It is often convenient to divide the 2^{d} RANS equation (1.13) by the mass density, to obtain an equation for the 'mean acceleration'

$$\boxed{\frac{DU_i}{Dt} \;:=\; \partial_t U_i \;+\; U_j\partial_{x_j}U_i \;=\; -\,\frac{1}{\rho}\partial_{x_i}P \;+\; \partial_{x_j}(2\nu S_{ij}(\mathbf{V})) \;-\; \partial_{x_j}\overline{u_i u_j}}\;. \tag{1.15}$$

We face a ***closure problem***: to close the system of the Reynolds equations (1.7) and (1.13), or, equivalently, of the equations (1.7) and (1.15), to be able to determine the first moments $P = \overline{p}$ and $U_i = \overline{v_i}$, we have to model the second moments

$$\overline{u_i u_j}\;.$$

A solution to this closure problem is a ***RANS model*** !..

1.2 Reynolds stress equations and models

1.2.1 Equations of the velocity fluctuations

By substracting, from the Navier-Stokes equation (1.8) divided by ρ, the RANS equation (1.15), we readily obtain the exact equation for the evolution of the velocity fluctuations

$$\frac{Du_i}{Dt} \;:=\; \partial_t u_i \;+\; U_j\partial_{x_j}u_i \;=\; -\,\frac{1}{\rho}\partial_{x_i}p' \;+\; \partial_{x_j}[\nu(\partial_{x_j}u_i) \;-\; u_i u_j \;+\; \overline{u_i u_j}] \;-\; u_j\partial_{x_j}U_i\;. \tag{1.16}$$

1.2.2 Equations of the second moments

Multiplying the equation (1.16) by u_j and averaging, then adding to it the mirror equation in which the indices i and j are interchanged, one gets, after some manipulations, exact equations for the evolution of the second moments, or the ***second-moments transport equations***,

$$\begin{aligned}\frac{D\overline{u_iu_j}}{Dt} \;:=\; \partial_t\overline{u_iu_j} \;+\; U_k\partial_{x_k}(\overline{u_iu_j}) \;=\; & \underbrace{-(\overline{u_iu_k}\;\partial_{x_k}U_j \;+\; \overline{u_ju_k}\;\partial_{x_k}U_i)}_{P_{ij}} \\ & \underbrace{-\frac{1}{\rho}\Big(\overline{u_i\;\partial_{x_j}p'} \;+\; \overline{u_j\;\partial_{x_i}p'}\Big)}_{\Pi_{ij}} \\ & -\underbrace{2\nu\overline{(\partial_{x_k}u_i)\;(\partial_{x_k}u_j)}}_{\varepsilon_{ij}} \\ & +\underbrace{\partial_{x_k}(\nu\partial_{x_k}\overline{u_iu_j})}_{D^\nu_{ij}} \;\underbrace{-\;\partial_{x_k}\overline{u_iu_ju_k}}_{D^t_{ij}}\;. \end{aligned} \tag{1.17}$$

By multiplication with $-\rho$, equivalent equations for the rate of change of the Reynolds stresses

$$\frac{D\tau_{ij}}{Dt} \;:=\; \partial_t\tau_{ij} \;+\; U_k\partial_{x_k}\tau_{ij} \;=\; \ldots \tag{1.18}$$

are readily obtained; those are the ***Reynolds-stress transport equations*** or more simply the ***Reynolds-stress equations***.
The equation (1.17) means that the rate of change of the second moment $\overline{u_iu_j}$ arises from an imbalance of the terms of the r.h.s.,

- the ***production terms*** P_{ij} , which are usually positive, hence their name;
- the ***pressure-velocity coupling terms*** Π_{ij} ;
- the ***dissipation terms*** ε_{ij} ;
- the ***viscous diffusion terms*** D^ν_{ij}, which are often negligible if the Reynolds number is high;
- the 3$^{\text{rd}}$ ***moments term*** D^t_{ij} .

We face a new ***closure problem***: the equations (1.17) for the 2$^{\text{d}}$ moments $\overline{u_iu_j}$ implies terms D^t_{ij} with the 3$^{\text{rd}}$ moments $\overline{u_iu_ju_k}$, that have to be modelled, as, also, Π_{ij} and ε_{ij} .
In ***stationary turbulence***, the rate of change $D\overline{u_iu_j}/Dt$ vanishes, hence the sum of the terms in the r.h.s. of the equation (1.17) vanish. Interestingly, these terms may be computed by sufficiently long DNS. Studying these terms and how they balance is relevant, and should help to model them. Such a study may be coined as a study of the ***second-moments budgets*** or ***Reynolds-stress budgets***.

1.2.3 Reynolds-stress models

Reynolds-stress models or ***second-moment models*** propose a model of Π_{ij} , ε_{ij} and D^t_{ij} , that yield a closed form of the equations (1.17). The coupled system of the equations (1.7), (1.15) and (1.17) allows then in principle the computation of the mean pressure P, mean velocity U_i and

second-moments $\overline{u_i u_j}$ i.e. Reynolds-stresses $\tau_{ij} = -\rho\overline{u_i u_j}$.
To develop Reynolds-stress models is a difficult task. An historical and seminal paper is the one of Launder *et al.* (1975). A more recent and pedagogical reference is the book Hanjalić & Launder (2011) which is almost entirely dedicated to this task, for free flows but also wall-bounded flows, which poses big problems.
From a computational point of view, the ***Reynolds-stress models*** are known to be, often, numerically unstable. They are also demanding in terms of computational power, since in 3D situations 6 equations (1.17) have to be solved, in addition to the mean mass conservation equation (1.7), and to the 3 mean momentum equations (1.15).

In this lecture, for the sake of simplicity, with an 'engineering' point of view, we will focus on simpler (but less accurate !) models, which rely on the ***eddy viscosity assumption*** that will be described now.

1.3 The eddy-viscosity assumption, turbulent kinetic energy and the corresponding momentum RANS equation

The idea undelying the ***eddy-viscosity assumption*** (or 'approximation') is attributed to Boussinesq in 1877, though then he examined a specific case. It is to write the Reynolds stresses τ_{ij} under a form similar to the one of the Cauchy stresses (0.3), i.e., after taking the Reynolds average,

$$\overline{\sigma_{ij}} \;=\; -\,P_{\text{static}}\delta_{ij} \;+\; 2\eta S_{ij}(\mathbf{V}) \;. \tag{1.19}$$

One writes thus

$$\tau_{ij} \;=\; -\,\alpha\delta_{ij} \;+\; 2\eta_t S_{ij}(\mathbf{V}) \;, \tag{1.20}$$

with α to be identified from the constraint that the trace of this tensor equation is strictly verified,

$$\tau_{ii} \;=\; -\,3\alpha \;, \tag{1.21}$$

and η_t the ***dynamic eddy*** or ***turbulent viscosity***. To write (1.20) is in general a crude assumption, as it will be shown in the problem 1.1, but this looks appealing, since it models the 6 unknown Reynolds stresses τ_{ij} with only 2 scalar variables, α and η_t .
By definition, see (1.14), the trace of the Reynolds stress tensor

$$\tau_{ii} \;=\; -\,2\rho k \tag{1.22}$$

with k the (specific) ***turbulent kinetic energy***

$$\boxed{k \;=\; \frac{1}{2}\overline{u_i u_i}} \;, \tag{1.23}$$

which has a clear physical meaning. Therefore the equation (1.21) gives

$$\alpha \;=\; \frac{2}{3}\rho k \;,$$

and consequently the ***eddy-viscosity assumption*** or ***Boussinesq assumption*** (1.20) reads

$$\boxed{\tau_{ij} \;=\; -\,\frac{2}{3}\rho k\delta_{ij} \;+\; 2\eta_t S_{ij}(\mathbf{V})} \quad\Longleftrightarrow\quad \boxed{\overline{u_i u_j} \;=\; \frac{2}{3}k\delta_{ij} \;-\; 2\nu_t S_{ij}(\mathbf{V})} \;, \tag{1.24}$$

with $\nu_t \;=\; \eta_t/\rho$ the ***kinematic eddy*** or ***turbulent viscosity***.

By injection into the ***momentum RANS equation*** (1.13), we get a new form of this equation,

$$\boxed{\rho[\partial_t U_i \ + \ \partial_{x_j}(U_i U_j)] \ = \ -\,\partial_{x_i}\Big(P \ + \ \frac{2}{3}\rho k\Big) \ + \ 2\partial_{x_j}\Big((\eta+\eta_t)S_{ij}(\mathbf{V})\Big)} \ . \tag{1.25}$$

- The term $\frac{2}{3}\rho k$ corresponds to a kind of ***'turbulent dynamic pressure'*** that increases the (modified) mean pressure P - compare with the formula that led to the Bernoulli theorems.
- The ***eddy viscosity*** η_t increases the material fluid viscosity η - this is coherent with the idea that ***turbulent mixing speeds up the diffusion***.

Dividing the equation (1.25) by ρ, we get the equivalent equation

$$\boxed{\frac{DU_i}{Dt} \ = \ -\,\partial_{x_i}\Big(\frac{P}{\rho} \ + \ \frac{2}{3}k\Big) \ + \ 2\partial_{x_j}\Big((\nu+\nu_t)S_{ij}(\mathbf{V})\Big)} \ . \tag{1.26}$$

1.4 The mixing length model of the eddy viscosity

The problem with equations (1.25) and (1.26) is mainly to estimate the ***eddy viscosity*** ν_t. Indeed it yields immediately the ***dynamic eddy viscosity***

$$\eta_t \ = \ \rho\nu_t \ . \tag{1.27}$$

The ***turbulent kinetic energy*** k is not needed in a simplistic but 'efficient' approach where one considers that it corresponds only to a technical correction[2] to the ***mean pressure*** P. A simple ***algebraic model***[3] of ν_t has been developed by Prandtl for ***shear flows*** where the ***mean strain-rate tensor***

$$S_{ij}(\mathbf{V}) = \frac{1}{2}\big[\partial_{x_j}(U_i) + \partial_{x_i}(U_j)\big] \ \neq \ 0$$

as measured by the ***mean strain-rate***

$$S \ = \ \sqrt{2S_{ij}S_{ij}} \ .$$

The idea is that dimensional analysis gives

$$\nu_t \ \equiv \ \ell \ v \ \equiv \ \ell^2 \ S$$

which is physically meaningful : the ***mean strain-rate*** S may be viewed as a source of mixing hence eddy viscosity. Hence Prandtl writes

$$\boxed{\nu_t \ = \ \ell_m^2 \ S} \tag{1.28}$$

where ℓ_m is the ***mixing length***. This length is estimated as the ***size of the largest turbulent eddies***, as it will be illustrated in problem 1.1. This model, useful because of its simplicity, is however not quite precise. Spalding and coworkers therefore developed the more sophisticated ***differential*** $k-\epsilon$ ***model***, that we will now present.

[2] This point of view is oversimplified: e.g., k influences the ***mixing*** and intervenes in many 'correlations' or closure formulas dealing with mixing or other 'complex' phenomena.

[3] 'Algebraic' refers to the fact that the eddy viscosity does not depend on fields other than U that are determined by differential equations.

1.5 The turbulent kinetic energy equation

1.5.1 The exact turbulent kinetic energy equation

By taking the trace of the ***second-moments transport equations*** (1.17), and dividing by 2, one gets the exact ***turbulent kinetic energy equation***

$$\frac{Dk}{Dt} \;:=\; \partial_t k \;+\; U_j \partial_{x_j} k \;=\; \underbrace{-\,\overline{u_i u_j}\;\partial_{x_j} U_i}_{P_k} \;\underbrace{-\,\frac{1}{\rho}\Big(\overline{u_i\;\partial_{x_i} p'}\Big)}_{\Pi} \;-\; \underbrace{\nu\overline{(\partial_{x_j} u_i)\;(\partial_{x_j} u_i)}}_{\varepsilon}$$
$$+\;\underbrace{\partial_{x_j}(\nu \partial_{x_j} k)}_{D^\nu} \;\underbrace{-\,\frac{1}{2}\partial_{x_j}\overline{u_i u_i u_j}}_{D^t} \;. \tag{1.29}$$

After a simple manipulation, that takes into account the fact that the fluctuating velocity field is conservative,

$$\partial_{x_i} u_i \;=\; 0 \;, \tag{1.30}$$

one may rewrite the ***turbulent kinetic energy equation*** (1.29) as follows,

$$\boxed{\frac{Dk}{Dt} \;=\; \underbrace{\partial_{x_j}(\nu \partial_{x_j} k)}_{D^\nu} \;\underbrace{-\,\partial_{x_j}\left(\overline{\frac{p' u_j}{\rho}} \;+\; \frac{1}{2}\overline{u_i u_i u_j}\right)}_{D_t} \;\underbrace{-\,\overline{u_i u_j}\;\partial_{x_j} U_i}_{P_k} \;-\; \underbrace{\nu\overline{(\partial_{x_j} u_i)\;(\partial_{x_j} u_i)}}_{\varepsilon}} \;. \tag{1.31}$$

In this equation, the different terms are typically the half-trace of terms in the ***second-moments transport equations*** (1.17), for instance

$$\varepsilon \text{ in equation (1.31)} \;=\; \frac{1}{2}\varepsilon_{ii} \text{ in equation (1.17)} \;.$$

1.5.2 The production term according to the eddy-viscosity assumption

The exact ***production term*** in equation (1.31),

$$\boxed{P_k \;=\; -\;\overline{u_i u_j}\;\partial_{x_j} U_i} \;, \tag{1.32}$$

may be rewrittent by the use of the eddy-viscosity assumption equation (1.24). This yields

$$P_k \;=\; 2\nu_t\; S_{ij}(\mathbf{V})\;\partial_{x_j} U_i \;=\; 2\nu_t\; S_{ij}(\mathbf{V})\; S_{ij}(\mathbf{V}) \;, \tag{1.33}$$

where we have used the fact that

$$\partial_{x_j} U_i \;=\; S_{ij}(\mathbf{V}) \;+\; \omega_{ij}(\mathbf{V}) \;,$$

where $\omega_{ij}(\mathbf{V})$, the rate-of-rotation tensor, is the antisymmetric part of the gradient of $\mathbf{V}$.
The last two expressions of P_k in the equations (1.33) clearly show that P_k is positive.

1.6 The turbulent dissipation rate

A sink, i.e., always negative term, in the equation (1.31) is the last one, $-\varepsilon$. It is the opposite of the (specific) ***turbulent (energy) dissipation rate***

$$\boxed{\varepsilon \;=\; \nu\;\overline{(\partial_{x_j} u_i)\;(\partial_{x_j} u_i)}} \;, \tag{1.34}$$

which is always positive.

An exact equation may be written for the time evolution of ε, see e.g. the section 3.4 of Hanjalić & Launder (2011). It is unfortunately quite complex. It has the form

$$\frac{D\varepsilon}{Dt} \;:=\; \partial_t\varepsilon \;+\; U_j\partial_{x_j}\varepsilon \;=\; \partial_{x_j}(\nu\partial_{x_j}\varepsilon) \;-\; 2\nu\overline{(\partial_{x_k}u_i)\,(\partial_{x_j}u_i)\,(\partial_{x_j}u_k)} \;+\; ... \tag{1.35}$$

where we have only explicited 2 of the 8 terms of the r.h.s. of the equation (3.15) of Hanjalić & Launder (2011), disregarding the volume force term[4].
It is interesting to analyze the dimensions of k and ε. Since a kinematic viscosity $\nu \;\equiv\; v\;\ell$ with v a velocity and ℓ a length, we get

$$k \;\equiv\; v^2 \quad \text{and} \quad \varepsilon \;\equiv\; v\;\ell\;(v/\ell)^2 \;\equiv\; v^3/\ell\;. \tag{1.36}$$

Consequently,

- k is a 'velocity-scale determining field', because from it we may define the velocity scale

$$v_k \;:=\; k^{1/2}\;; \tag{1.37}$$

- ε is a 'length-scale determining field', because from it and using (1.37) we may define the length scale

$$\ell_{k\varepsilon} \;:=\; \frac{v_k^3}{\varepsilon} \;=\; \frac{k^{3/2}}{\varepsilon}\;. \tag{1.38}$$

It is important to state that ε is often close to the ***classical turbulent dissipation rate***

$$\boxed{\epsilon \;=\; 2\nu\;\overline{S_{ij}(\mathbf{u})\;S_{ij}(\mathbf{u})}}\;. \tag{1.39}$$

Indeed, one may show that

$$\epsilon \;-\; \varepsilon \;=\; \nu\;\partial_{x_j}\overline{u_i\partial_{x_i}u_j}\;, \tag{1.40}$$

which vanishes if the turbulence is ***'homogenous'*** i.e. uniform in space, i.e., if all mean properties of turbulence do not depend on the position. This is important because

$$\varepsilon \;\simeq\; \epsilon \tag{1.41}$$

appears as the ***specific power injected in the 'Kolmogorov cascade'***, see e.g. Plaut (2021*a*).

1.7 The $k-\varepsilon$ model

1.7.1 The main idea - The eddy-viscosity formula

The $k-\varepsilon$ model was designed by Spalding and co-workers at Imperial College in the late 1960s - early 1970s[5], and presented in Launder & Spalding (1974). The main idea is that k and ε are the most relevant 'turbulent fields' that may characterize the local properties of turbulence and the turbulent viscosity. By dimensional analysis, one assumes that

$$\nu_t \;=\; C_\nu\;v_k\;\ell_{k\varepsilon}\;, \tag{1.42}$$

[4] See the last remark of the chapter 0.

[5] The history of this model is sketched in the section 7.4.2, 'Popular two-equation eddy-viscosity models' of Hanjalić & Launder (2011).

with C_ν (or C_μ) a dimensionless number. From the equations (1.37) and (1.38), we get

$$\boxed{\nu_t \;=\; C_\nu \,\frac{k^2}{\varepsilon}}\;, \tag{1.43}$$

which displays reasonable physical tendencies:

- if the turbulent kinetic energy k increases, at fixed ε one expects a more efficient turbulent mixing, and ν_t does increase ;
- if the turbulent dissipation rate ε increases, at fixed k, one expects a less efficient turbulent mixing, and ν_t does decrease.

1.7.2 The modeled turbulent kinetic energy equation

With the idea that the term D_t in the exact turbulent kinetic energy equation (1.31) corresponds essentially to turbulent diffusion effects, one writes it using the turbulent viscosity. This leads to the ***modeled turbulent kinetic energy equation***

$$\boxed{\frac{Dk}{Dt} \;=\; \underbrace{\partial_{x_j}\left(\left(\nu+\frac{\nu_t}{\sigma_k}\right)\partial_{x_j}k\right)}_{D^\nu+D^t} \;+\; P_k \;-\; \varepsilon}\;, \tag{1.44}$$

with σ_k a dimensionless number of order 1, and the production term P_k already studied in section 1.5.2.

1.7.3 The modeled turbulent dissipation equation

Contrarily to what has been done with k, because the exact ε equation (1.35) is quite complex, it is not relevant to use this equation as a a starting point to obtain a modeled equation for ε. Instead, Spalding and coworkers wrote an equation analogous in form to the modeled equation (1.44) for k, with viscous diffusion, turbulent diffusion, production and dissipation terms that have the correct dimension by multiplication by the factor ε/k for the last two terms, inserting also numbers to have new degrees of freedom for good modeling:

$$\boxed{\frac{D\varepsilon}{Dt} \;=\; \underbrace{\partial_{x_j}\left(\left(\nu+\frac{\nu_t}{\sigma_\varepsilon}\right)\partial_{x_j}\varepsilon\right)}_{D_\varepsilon^\nu+D_\varepsilon^t} \;+\; C_1\,\frac{\varepsilon}{k}\,P_k \;-\; C_2\,\frac{\varepsilon}{k}\,\varepsilon}\;. \tag{1.45}$$

There σ_ε , C_1 and C_2 are dimensionless numbers of order 1.

1.7.4 Coefficients

The ***coefficients of the*** $k-\varepsilon$ ***model*** proposed by Launder & Spalding (1974), on the basis of the study of representative flow cases that were at this time documented experimentally, are

$$\boxed{C_\nu \;=\; 0.09\;,\quad \sigma_k \;=\; 1\;,\quad \sigma_\varepsilon \;=\; 1.3\;,\quad C_1 \;=\; 1.44\;,\quad C_2 \;=\; 1.92}\;. \tag{1.46}$$

1.7.5 Problem

We follow in this problem some processes imagined by Spalding and coworkers to set up or 'verify' the $k-\varepsilon$ model, and 'fix' or 'fit' some of its constants. We illustrate their ideas by the use of the DNS database (1.1), and thus go beyond their work of the 1960s - 1970s. Indeed, we have more knowledge now: DNS of turbulent channel flow are numerous and well documented, whereas when the $k-\varepsilon$ model was invented they were just infeasible.
This problem is of course also an occasion to gain some general knowledge on turbulence phenomena and on the RANS approach.

Problem 1.1 ***Turbulent channel flow: RANS mixing length and $k-\varepsilon$ models***

We consider, as a prototype of a well documented ***wall-bounded turbulent flow***, a ***turbulent channel flow*** as sketched on the figure 1.1a. The half-channel height is δ, x is the streamwise coordinate, y the wall-normal coordinate, z the spanwise coordinate. In the region of interest, far from the inlet and outlet, and not too close to the sidewalls at $z = \pm L_z/2$, the ***mean velocity field***

$$\boxed{\mathbf{V} \;=\; U(y)\;\mathbf{e}_x} \tag{1.47}$$

and the ***mean*** (modified in the sense of section 0.2) ***pressure field***

$$P \;=\; -\,Gx \;+\; p_0(y) \quad \text{with} \quad G>0 \text{ the } \textbf{\textit{opposite of the mean pressure gradient}}. \tag{1.48}$$

All the other mean fields depend only on y. In part 1, after some generalities, we zoom on the near-wall region, and use the simple historical 'algebraic' ***mixing-length RANS model*** of section 1.4 that was developed prior to the $k-\varepsilon$ model, to evidence the three-layers structure depicted on the figure 1.1b. In parts 2 and 3, this structure is evidenced and further characterized with the DNS database (1.1).

Part 1 ***General study, within the RANS framework, of the mean-flow profile***

1.1 Explicit the 3 components of the ***RANS momentum equation*** (1.13).

- With the x-component, show that there exists a balance between the mean pressure gradient, viscous and turbulence effects.
- With the y-component, show that, as one goes off the wall, the pressure tends to diminish because of the turbulent fluctuations.
- With the z-component, show that the Reynolds stress τ_{yz} vanishes, which may be viewed as a 'symmetry' property.

1.2 Integrate the x-component of the ***RANS momentum equation*** with respect to y, to show that

$$\eta\frac{dU}{dy} \;+\; \tau_{xy} \;=\; \tau_w \;-\; Gy\;, \tag{1.49}$$

with the mean viscous stress, or 'friction', at the wall, or ***mean wall shear stress***

$$\tau_w \;=\; 2\eta S_{xy}(\mathbf{V})|_{y=0} \;=\; \eta\frac{dU}{dy}\Big|_{y=0} \;=\; \eta U_0'\;. \tag{1.50}$$

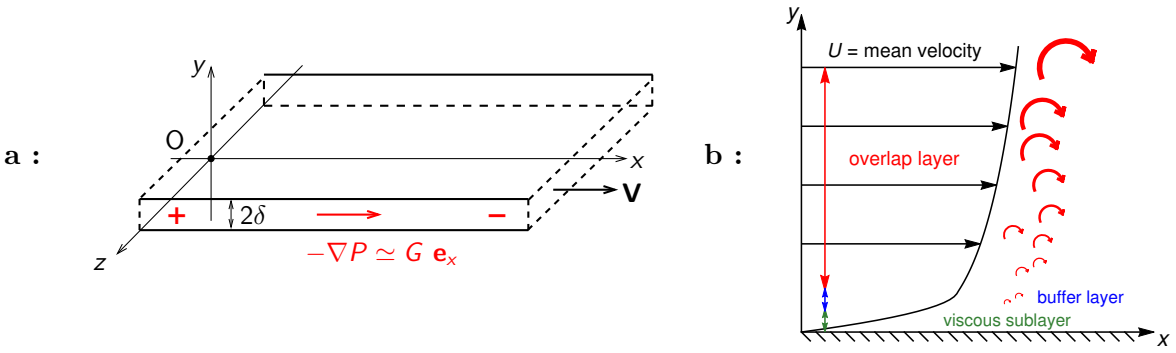

Fig. 1.1 : (color online) **a:** Sketch of a ***turbulent channel flow*** setup as studied in the problem 1.1. **b:** Zoom on the region near the lower wall, that is somehow a ***turbulent boundary layer***. The curved arrows represent turbulent eddies.

This enables the definition of ***'wall units'*** of velocity, using as the velocity scale the ***friction velocity***

$$\boxed{u_\tau \;=\; \sqrt{\tau_w/\rho}} \quad\Longleftrightarrow\quad \boxed{\tau_w \;=\; \rho u_\tau^2}\;, \tag{1.51}$$

and of length, using as the length scale the ***friction length***

$$\boxed{\ell_\tau \;=\; \nu/u_\tau}\;. \tag{1.52}$$

Thus we define

$$\boxed{U^+ \;=\; \frac{U}{u_\tau}} \quad\text{and}\quad \boxed{y^+ \;=\; \frac{y}{\ell_\tau} \;=\; \frac{y u_\tau}{\nu}}\;. \tag{1.53}$$

1.3 One observes in experiments and DNS the ***mirror symmetry with respect to the center-plane*** at $y=\delta$,

$$U(2\delta - y) \;=\; U(y) \quad\text{and}\quad \tau_{xy}(2\delta - y) \;=\; -\,\tau_{xy}(y)\;. \tag{1.54}$$

Deduce from this the symmetry properties of the velocity gradient

$$\frac{dU}{dy}(2\delta - y) \;=\; \qquad\Longrightarrow\quad \frac{dU}{dy}(y=\delta) \;=\; \quad, \tag{1.55}$$

and determine also the main Reynolds stress at the centerplane

$$\tau_{xy}(y=\delta) \;=\; \quad. \tag{1.56}$$

1.4 Deduce from this and the (integrated) ***RANS momentum equation*** (1.49) the expression of the mean wall shear stress as a function of the mean pressure gradient and the half-channel height,

$$\tau_w \;=\; \quad. \tag{1.57}$$

1.5 Find this same relation with a global mean momentum budget of the fluid in a region of interest, in the x direction, and explain the physics behind this.

1.6 Show that there exists a ***viscous sublayer***, for y quite small, where the profile of mean velocity is linear in y,

$$U(y) \simeq U_0' \, y \quad \Longleftrightarrow \quad \boxed{U^+(y^+) \simeq y^+} \, . \tag{1.58}$$

1.7 Show that, farer from the wall, in the ***overlap layer*** where both

- y is sufficiently large, such that the Reynolds stress τ_{xy} dominates the viscous stress $\eta dU/dy$

and

- $y \ll \delta$,

one may assume that the Reynolds stress is roughly constant,

$$\tau_{xy} \simeq \qquad . \tag{1.59}$$

1.8 Following Boussinesq, Prandtl and von Karman, we use in this overlap layer an eddy-viscosity model to calculate the Reynolds stress τ_{xy} , with a ***mixing-length model*** for the eddy-viscosity

$$\boxed{\eta_t = \rho \nu_t = \rho \, \frac{dU}{dy} \, \ell_m^2} \, , \quad \boxed{\ell_m = \kappa \, y} \, , \quad \boxed{\kappa = \textbf{\textit{von Karman constant}}} \, . \tag{1.60}$$

Show that this implies that the ***overlap layer*** is a ***log layer*** where the ***log law*** holds,

$$\boxed{U^+(y^+) \simeq \frac{1}{\kappa} \ln y^+ + C} \tag{1.61}$$

with C an additive constant.

1.9 Calculate the eddy viscosity in the ***log layer***, and show that it has a simple y-dependence

$$\eta_t = \qquad \Longleftrightarrow \quad \boxed{\nu^+ := \frac{\eta_t}{\eta} = \qquad} \, . \tag{1.62}$$

Part 2 ***Evidence and characterization of the viscous sublayer on the DNS database***

2 Open with a web browser the database (1.1), and save in a folder the files

`LM_Channel_0180_mean_prof.dat`

and

`LM_Channel_5200_mean_prof.dat`

that give data about the mean flow profile for two DNS at two different ***friction Reynolds numbers***

$$\boxed{Re_\tau \;=\; \frac{u_\tau \delta}{\nu}}\;, \tag{1.63}$$

$Re_\tau \;=\; 180$ and 5200.

With a **Matlab** program that looks like this (the ... should be corrected !)

```
data= load("LM_Channel_0180_mean_prof.dat");
%% Extraction of the column 2 : tabulated values of y+
yp= data(:,2);
%% Extraction of the column ... : tabulated values of U+
Up= data(:,...);
%% Figure: U+ function of y+, DNS and viscous law
figure(1); hold on;
plot( yp, Up, "k","LineWidth",2); xlabel("y+"); ylabel("U+");
plot( yp, yp, "g","LineWidth",1);
```

compare the DNS mean-flow profiles to the ***viscous sublayer*** profile given by the equation (1.58), i.e., realize the plot asked for in the figure 1.2a. With the criterion that

$$\text{diff} \;=\; |U^+ \;-\; y^+| \;=\; y^+ \;-\; U^+ \;<\; 0.1\;, \tag{1.64}$$

by plotting this difference, in figure 1.2b, define with one digit the ***upper boundary of the viscous sublayer***,

$$y_0^+ \;\simeq\; \quad . \tag{1.65}$$

Check that, when the friction Reynolds number is increased, the profiles look similar and y_0^+ computed with one digit is unchanged.

a : **b :**

Fig. 1.2 : DIY ! Evidence and characterization of the ***viscous sublayer*** on the ***turbulent channel flow DNS*** of the database (1.1) at $Re_\tau = 180$. **a :** The profiles of U^+ (thick continuous curve) and the viscous law $U^+ = y^+$ (thin dashed line). **b :** The difference (1.64) (thick line and data points) and the limit 0.1 (dashed line).

Part 3 ***Evidence and characterization of the log layer on the DNS database***

3.1 With a **Matlab** program, for the DNS of the database at $Re_\tau = 5200$, compute the first indicator function of the log law (1.61),

$$\alpha(y^+) \ := \ y^+ \ S^+ \tag{1.66}$$

with the ***mean strain rate***

$$S^+ \ := \ \frac{dU^+}{dy^+} \ . \tag{1.67}$$

Plot $\alpha(y^+)$ vs y^+, i.e., realize the plot asked for in the figure 1.3a. Check that $\alpha(y^+)$ displays a kind of plateau around an inflexion point, and deduce from the position of this point estimates, with one digit, of y^+ in the core of the ***log layer*** and, with two digits, of the ***von Karman constant***,

$$y_1^+ \ \simeq \qquad\qquad \text{and} \qquad \boxed{\kappa \ \simeq \qquad} \ . \tag{1.68}$$

Check that κ is close to the value that was admitted in the 1970s, and used by Spalding and coworkers,

$$\boxed{\kappa \ = \ 0.43} \ . \tag{1.69}$$

3.2 For the same DNS, plot also the second indicator function of the log law (1.61),

$$\gamma(y^+) \ := \ U^+(y^+) \ - \ \kappa^{-1} \ln y^+ \ , \tag{1.70}$$

by using *your* measured value of κ (1.68), i.e., realize the plot asked for in the figure 1.3b. Check that $\gamma(y^+)$ does not vary too much around y_1^+ determined in (1.68), and estimate the additive constant C in the log law (1.61) with two digits,

$$C \ = \qquad . \tag{1.71}$$

3.3 For the same DNS, realize the figure 1.3c as defined in the caption.
Conclude regarding these models of the mean flow.

Comments on parts 1 to 3 ***Universality of this three-layers structure***

The ***three-layers structure*** evidenced in parts 1 to 3, with a ***viscous sublayer***, a ***buffer layer*** and finally a ***log layer***, depicted on the figure 1.1b, and more precisely displayed on the figure 1.3c, has been confirmed recently in turbulent channel flows by DNS at even higher Reynolds number, see the figure 2 of Yamamoto & Tsuji (2018). This three-layers structure 'exists' in ***many other wall-bounded turbulent flows*** than channel flows. The commas around 'exists' refer to the log layer, which 'exists' not exactly, but only 'up to a certain precision'...

- In ***turbulent zero-pressure gradient boundary layers over a flat plate***, the analysis of part 1 may be redone. Locally, neglecting the x-dependence, and close enough to the wall, one obtains an equation similar to (1.49) but with $G = 0$, i.e., after division by ρ,

$$\nu \frac{dU}{dy} \; - \; \overline{u_x u_y} \; \simeq \; u_\tau^2 \; . \tag{1.72}$$

  A numerical confirmation of the log law is for instance presented in the figure 4 of Eitel-Amor *et al.* (2014). By the way, some (e.g. the meteorologists that study 'atmospheric boundary layers') name in equation (1.72) the 1st term of the l.h.s. the (kinematic) viscous momentum flux, the 2d term of the l.h.s. the (kinematic) turbulent momentum flux, and say that the part of the boundary layer where their sum is indeed 'constant' is a ***'constant-flux layer'***.

- In ***turbulent pipe flows***, the analysis of part 1 may be redone; one has to use cylindrical coordinates, there is also a mean pressure gradient G... finally the same log law is found. A numerical confirmation is for instance presented in the figure 6 of El Khoury *et al.* (2013).

Therefore, one may speak of a ***universal structure***, and one may 'believe' that this three-layers structure 'exists' in almost all wall-bounded flows, though the precise values of the von Karman constant κ and of the additive constant C may depend on the flow case. This ***universality*** pushed Spalding et al. to verify their $k-\varepsilon$ model in 'the' log layer, and to use this flow case to adjust two of the model coefficients, i.e. C_ν and σ_ε . The hope was also to obtain near-wall 'boundary conditions' for k and ε in the log layer, see the equations (1.73). The study of the log layer with the $k-\varepsilon$ model is the object of the part 4 of our problem.

Part 4 ***Analytical study of the $k-\varepsilon$ model in the log layer***

4.1 In the ***overlap*** or ***log layer***, we assume that the $k-\varepsilon$ model is as relevant as the mixing-length model introduced in (1.60). With the idea that the diffusion terms in the equation for k are negligible in front of the production and dissipation terms, i.e., that there is a ***production - dissipation balance***[6], determine k and ε, i.e., establish ***wall laws*** for k and ε,

$$k \; = \qquad\qquad , \quad \varepsilon \; = \qquad\qquad . \tag{1.73}$$

Comment.

4.2 Show that, according to the $k-\varepsilon$ model, in the ***log layer***, one has the ***Townsend's relation***[7]

$$\frac{\nu_t \; dU/dy}{k} \; = \; - \frac{\overline{u_x u_y}}{k} \; = \; \text{constant} \; = \tag{1.74}$$

which depends simply of one coefficient of the model. Comment.

[6] Townsend, a british physicist active in the middle of the XXth century, therefore designates a log layer as an ***'equilibrium layer'***.

[7] See Townsend (1961). Some, however, attribute the relation (1.74) to Bradshaw *et al.* (1967).

a : **b :**

c :

Fig. 1.3 : DIY ! Evidence and characterization of the ***log layer*** on the ***turbulent channel flow DNS*** of the database (1.1) at $Re_\tau = 5200$. **a :** The first indicator function α (1.66). **b :** The second indicator function γ (1.70). **c :** The DNS mean flow (continuous curve) with the viscous layer profile for $y^+ < 10$ and the log law profile for $y^+ > 10$ (dashed curves). In all plots, the vertical dashed line shows $y^+ = y_1^+$ which locates the log layer, see the first of the equations (1.68).

4.3 Show that the (modeled) equation (1.45) of the turbulent dissipation is valid in the ***log layer***, provided that there is a relation between the von Karman constant κ and some coefficients of the model $k - \varepsilon$:

$$\kappa^2 \;=\; \qquad . \tag{1.75}$$

Verify this with the coefficients (1.46), using the 'historical' value (1.69) of κ .

We move now to a study, based on the database (1.1), of the ***second moments of the velocities*** (equivalently, the ***Reynolds stresses***), ***turbulent kinetic energy***, ***eddy viscosity*** and ***turbulent dissipation***, and to detailed comparisons with the $k - \varepsilon$ ***model*** and ***wall laws***.

Part 5 ***Study of the second moments of the velocities, the turbulent kinetic energy and the eddy viscosity on the DNS database***

5.1 Download from the database (1.1) the file `LM_Channel_0180_vel_fluc_prof.dat` that give the ***second moments of the velocities*** for the DNS at $Re_\tau = 180$. Wall units are used: the 3^{rd} column for instance, that the authors denote $u'u'$, gives with our notations the dimensionless ***variance***

$$\boxed{\overline{u_x^+ u_x^+} \;=\; \frac{\overline{u_x u_x}}{u_\tau^2} \;=\; \frac{\rho\overline{u_x u_x}}{\rho u_\tau^2} \;=\; -\,\tau_{xx}^+} \,. \tag{1.76}$$

Extract from this file $\overline{u_x^+ u_x^+}$, $\overline{u_y^+ u_y^+}$, $\overline{u_z^+ u_z^+}$ and the dimensionless ***turbulent kinetic energy***

$$\boxed{k^+ \;=\; \frac{k}{u_\tau^2}} \,. \tag{1.77}$$

Plot these quantities vs y/δ with continuous lines, adding on the three first plots the prediction of the ***ideal***[8] $k-\varepsilon$ ***model*** with dashed lines, i.e., according to the eddy-viscosity assumption (1.24),

$$\overline{u_x^+ u_x^+} \;=\; \overline{u_y^+ u_y^+} \;=\; \overline{u_z^+ u_z^+} \;=\; \frac{2}{3}k^+ \;, \tag{1.78}$$

i.e., realize the plots asked for in the figures 1.4abcd. Comment in terms of ***'isotropy'*** or ***'anisotropy'*** of the variances, and discuss the influence of the wall at $y=0$. In particular, evidence a ***wall-damping effect*** on two variances, a ***wall-enhancing effect*** on one variance, that are not captured by the ideal $k-\varepsilon$ model. Is this issue specific to the $k-\varepsilon$ model, or could it be more general ?

5.2 Check that, in terms of the dimensionless ***eddy viscosity*** ν^+ defined in (1.62), the eddy-viscosity assumption (1.24) yields dimensionless ***covariances*** of the velocities, for $i \neq j$,

$$\begin{array}{rcll} \overline{u_i^+ u_j^+} & = & -\nu^+ \dfrac{dU^+}{dy^+} & \text{if} \qquad (i,j) = (x,y) \text{ or } (y,x) \;, \\ \overline{u_i^+ u_j^+} & = & 0 & \text{otherwise.} \end{array} \tag{1.79}$$

5.3 Verify with **Matlab** that, for the DNS at $Re_\tau = 180$, the maximal absolute values of the[9] ***covariances*** that should vanish are indeed small as compared with the maximal absolute value of $\overline{u_x^+ u_y^+}$,

$$\max\left(\max\left|\overline{u_x^+ u_z^+}\right| = \quad , \; \max\left|\overline{u_y^+ u_z^+}\right| = \quad\right) \ll \left(\max\left|\overline{u_x^+ u_y^+}\right| = \quad\right). \tag{1.80}$$

Comment.

5.4 Realize the plots asked for in figure 1.4e of the opposite of the ***main covariance*** $\overline{u_x^+ u_y^+}$ and the ***mean strain rate*** S^+ (1.67). Deduce from them the ***'exact' eddy viscosity***

$$\nu^+ \;=\; -\frac{\overline{u_x^+ u_y^+}}{S^+} \;, \tag{1.81}$$

and plot it in figure 1.4f with its log-layer model $\nu^+ = \kappa y^+$, see the equations (1.62). For this purpose, explicit generally the relation between y^+ and y/δ,

$$y^+ \;=\; \quad . \tag{1.82}$$

Comment this relation and the figures 1.4ef.

Indications:

- The file `LM_Channel_0180_mean_prof.dat` already used in part 2 is necessary to get the values of $S^+ = dU^+/dy^+$.
- Use `./` to divide term by term two vectors.
- Use the value of κ measured in (1.68).

5.5 With the files of the DNS at the highest value of Re_τ accessible in the database, $Re_\tau = 5200$, realize the plots asked for in figure 1.4g.
Compare with the figure 1.4f and comment.

[8] This would be the ***'ideal'*** $k-\varepsilon$ ***model*** since it would predict a profile of $k^+(y)$ strictly equal to the one found naturally in the DNS. This cannot be !.. Thus the mismatchs between the velocity variances and the 'ideal' $k-\varepsilon$ model predictions visible on the figures 1.4abc get, typically, worse, if the 'real' $k-\varepsilon$ model is used.

[9] From now on we do not recall, in the text of the problem 1.1, that all fields are dimensionless.

a :

b :

c :

d :

e :

f :

g :

Fig. 1.4 : DIY ! Various turbulent fields from the DNS database (1.1), in dimensionless wall units, plotted vs y/δ. In figures a to f, $Re_\tau = 180$, whereas, in figure g, $Re_\tau = 5200$. **a :** the continuous line shows the ***variance*** $\overline{u_x^+ u_x^+}$, see (1.76), the dashed line the profile according to the eddy-viscosity assumption, i.e., $\frac{2}{3}k^+$, see (1.78). **b :** the same for the ***variance*** $\overline{u_y^+ u_y^+}$. **c :** the same for the ***variance*** $\overline{u_z^+ u_z^+}$. **d :** the ***turbulent kinetic energy*** k^+ (1.77). **e :** the continuous line shows the opposite of the ***main covariance*** $\overline{u_x^+ u_y^+}$, the dashed line the ***mean strain rate*** $S^+ = dU^+/dy^+$. **f :** the continuous line shows the ***'exact' eddy viscosity*** $\nu^+ = -\overline{u_x^+ u_y^+}/S^+$, the dashed line its log-layer model $\nu^+ = \kappa y^+$. **g :** Same as figure f but for $Re_\tau = 5200$. The vertical dashed line shows $y^+ = y_1^+$ which locates the log layer, see the first of the equations (1.68).

Part 6 ***Study of the turbulent kinetic energy, of its production, and of the turbulent dissipation on the DNS database***

We perform a systematic study of these fields, on the DNS database (1.1), for typical values of the friction Reynolds number, $Re_\tau = 180,\ 1000$ and 5200.

6.1 For these DNS, realize the plots asked for in figures 1.5ab, of k^+ vs y/δ in lin-lin scales, vs y^+ in log-lin scales, with lines of increasing thicknesses and different colors. The range of values of y^+ for the figure 1.5b (and also dfh) should be large, i.e., one should start deep in the viscous sublayer, at $y^+ = 0.1$, and reach at the end the maximum value of y^+. Comment, by identifying a near-wall peak of k^+ at a 'robust' value of y^+, with 1 digit,

$$y^+_{\max k} \ \simeq \qquad . \tag{1.83}$$

Indication: for log-lin scales, use the command `semilogx` instead of `plot`.

6.2 Do these DNS confirm the wall law for k (1.73) in the log layer, that reads in wall units

$$k^+ \ = \ \frac{1}{\sqrt{C_\nu}} \quad ? \tag{1.84}$$

6.3 Download from the database (1.1) the files `LM_Channel_*_RSTE_k_prof.dat` that give data about the ***exact turbulent kinetic energy equation*** (1.31), for the same DNS. Remark that one column contains the ***production term*** (1.32), here, because of the form of the mean flow

$$P_k \ = \ -\,\overline{u_x u_y}\,\frac{dU}{dy}\ , \tag{1.85}$$

in wall units, i.e.

$$\boxed{P_k^+ \ = \ \frac{\ell_\tau}{u_\tau^3} P_k \ = \ \frac{\nu}{u_\tau^4} P_k \ = \ -\,\overline{u_x^+ u_y^+}\,\frac{dU^+}{dy^+} \ = \ -\,\overline{u_x^+ u_y^+}\ S^+} \tag{1.86}$$

with the notation (1.67). Realize the plots asked for in figures 1.5cd of P_k^+ vs y/δ or y^+. Comment, by identifying a near-wall peak of P_k^+ at a 'robust' value of y^+, with 2 digits,

$$y^+_{\max P_k} \ \simeq \qquad . \tag{1.87}$$

6.4 Accordingly, identify in the files `LM_Channel_*_RSTE_k_prof.dat` the column that gives the ***turbulent dissipation rate*** in wall units,

$$\boxed{\varepsilon^+ \ = \ \frac{\nu}{u_\tau^4}\varepsilon}\ . \tag{1.88}$$

Realize the plots asked for in figures 1.5ef of ε^+ vs y/δ or y^+. Comment.

6.5 Do these DNS confirm the idea that there would exist a universal value of ε^+ at the wall ?

6.6 To test the wall law for ε (1.73) in the log layer, transform it to wall units

$$\varepsilon^+ \ = \qquad \Longleftrightarrow \quad y^+\varepsilon^+ \ = \qquad . \tag{1.89}$$

Realize the plots asked for in figures 1.5gh. Comment.

Indication: use `.*` to multiply term by term two vectors.

a : **b :**

c : **d :**

e : **f :**

g : **h :**

Fig. 1.5 : DIY ! Various turbulent fields extracted from the DNS database (1.1), in dimensionless wall units, at $Re_\tau = 180$ (thin black line), 1000 (thicker blue line) and 5200 (thickest red line), plotted vs y/δ in lin-lin scales (left) or y^+ in log-lin scales (right). **a,b :** ***turbulent kinetic energy*** k^+ (1.77); the vertical dashed line on figure b indicates $y^+_{\max k}$ defined in the equation (1.83). **c,d :** ***production term*** P_k^+ (1.86); the vertical dashed line on figure d indicates $y^+_{\max P_k}$ defined in the equation (1.87). **e,f :** ***turbulent dissipation rate*** ε^+ (1.88). **g,h :** ***premultiplied dissipation rate*** $y^+\varepsilon^+$.

Comment on one particular result of part 6 - question 6.3

The figure 1.5d suggests that the production term in wall units $P_k^+ = -\overline{u_x^+ u_y^+}\ S^+$ converges to a ***universal asymptotic profile*** as $Re_\tau \to +\infty$. Indeed a ***general asymptotic theory*** that yields asymptotic expressions of the main covariance $-\overline{u_x^+ u_y^+}$, the mean strain rate S^+ and, consequently, P_k and ν_t , valid for channel flows and also boundary layer and pipe flows as $Re_\tau \to +\infty$, has been recently presented in Heinz (2018, 2019), see the problem 1.3 at the end of this chapter.

Part 7 ***Study of the eddy viscosity, of the ratio k^2/ε and of the eddy-viscosity coefficient C_ν on the DNS database***

We continue our systematic study of the DNS database (1.1), for typical values of the friction Reynolds number, $Re_\tau = 180,\ 1000$ and 5200.

7.1 Precise the dimensionless form of the $k-\varepsilon$ ***model eddy-viscosity law*** (1.43), in terms of $\nu^+ = \eta_t/\eta = \nu_t/\nu$, k^+ and ε^+,

$$\boxed{\nu^+ \ =} \quad . \tag{1.90}$$

7.2 With a **Matlab** program similar to the one used in question 10.3, plot the ***'exact' eddy viscosity*** $\nu^+ = -\overline{u_x^+ u_y^+}/S^+$ (1.81) vs y/δ or y^+ for the 3 DNS of focus, i.e. realize the plots asked for in the figures 1.6ab. The range of values of y^+ for the figures 1.6bdfh should be the same as the one used in the figures 1.5bdfh. Comment.

7.3 On your screen, or using the `axis` command, zoom on the range $y^+ \in [0.1, 10]$, and check that, in the viscous sublayer, as defined by (1.64, 1.65), the eddy viscosity

$$\nu^+ \ < \ 0.1 \tag{1.91}$$

for all DNS. Explain what this means physically.

7.4 Extend your program to compute k^{+2}/ε^+ for the DNS already studied, and realize the plots asked for in the figures 1.6cd. Comment.

Indication: use `.^2` to raise all terms of a vector to the power 2, `./` to divide term by term two vectors.

7.5 Extend your program to compute 'ideal' functions $c_\nu(y^+; Re_\tau)$ that would be defined by the fact that the $k-\varepsilon$ ***model eddy-viscosity law*** (1.90) is exactly valid by the use of c_ν instead of C_ν , i.e.

$$c_\nu \ := \ \frac{\nu^+}{k^{+2}/\varepsilon^+} \ . \tag{1.92}$$

Realize the plots asked for in the figures 1.6ef. Comment briefly.

7.6 Compute the ratio of the exact DNS eddy viscosity to the $k-\varepsilon$ model eddy-viscosity,

$$R_\nu \ := \ \frac{\nu^+}{0.09\ k^{+2}/\varepsilon^+} \ , \tag{1.93}$$

and realize the plots asked for in the figures 1.6gh.
Given your answer to the question 7.3, is the divergence visible in the region $y^+ < 1$ on the figures 1.6fh important ?
Evidence a ***wall-damping effect*** on the ***eddy viscosity***.

a : **b :**

c : **d :**

e : **f :**

g : **h :**

Fig. 1.6 : DIY ! Various turbulent fields extracted from the DNS database (1.1), in dimensionless wall units, at $Re_\tau = 180$ (thin black line), 1000 (thicker blue line) and 5200 (thickest red line), plotted vs y/δ in lin-lin scales (left) or y^+ in log-lin scales (right). **a,b :** ***'exact' eddy viscosity*** ν^+. **c,d :** ratio k^{+2}/ε^+. **e,f :** ***ideal eddy viscosity 'function'*** c_ν defined by the equation (1.92). **g,h :** ***eddy-viscosity ratio*** R_ν defined by the equation (1.93). In figures fh, the vertical dashed line shows $y^+ = y_0^+$, the upper boundary of the viscous sublayer, see the equation (1.65).

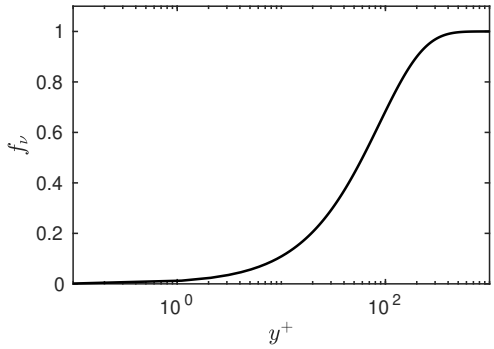

Fig. 1.7 : The ***eddy-viscosity damping function*** f_ν (1.94) of Chien (1982) vs y^+. The facts that $f_\nu(y^+) \longrightarrow 0$ as $y^+ \to 0$, $f_\nu(y^+) \longrightarrow 1$ as $y^+ \to +\infty$ are typical of a damping function.

1.7.6 Some conclusions regarding the $k-\varepsilon$ model

Our answers to the question 5.1 of the problem 1.1, see the plots of the velocity variances $\overline{u_x^+ u_x^+}$, $\overline{u_y^+ u_y^+}$ and $\overline{u_z^+ u_z^+}$ of the figures 1.4abc, show that all models that use the ***eddy-viscosity-assumption*** cannot describe the ***wall-damping*** and ***wall-enhancing effects*** that lead to an ***anisotropy of the velocity variances near a solid wall***. A solution to this issue may only be obtained with ***second moments*** or ***Reynolds stress models***, which are beyond the scope of this lecture.

Disregarding this issue, with the idea that, for engineers, a good knowledge of the turbulent kinetic energy k might be sufficient to characterize roughly the turbulent fluctuations, there remains obviously the issue of the choice of the ***boundary conditions near a solid wall***.
In the ***'high-Reynolds number'*** version of the $k-\varepsilon$ model, one does not mesh the viscous sublayer, but starts computations somewhere around[10] $y^+ \simeq 100$, at a point which stands hopefully in the ***log layer***, and uses the ***wall laws*** (1.61) for the mean flow U, (1.73) for the turbulent kinetic energy k and the turbulent dissipation ε. A complexity is that the value of the wall shear stress is needed, which may be estimated roughly and then refined. Anyhow, our answers to the questions 6 of the problem 1.1, see the figures 1.5b of k^+ and 1.5h of $y^+\varepsilon^+$, clearly show that these boundary conditions are poorly relevant: this leads typically to inaccurate results.
This explains why, very soon, the developpers of the $k-\varepsilon$ model, and other authors, proposed ***'low-Reynolds number'*** versions of this model, that require a meshing down to the wall, and rely on modifications of some coefficients of the model that become ***'damping functions'*** of the wall distance or of other fields. For instance, in Chien (1982), a damping function[11]

$$f_\nu \;=\; f_\nu(y^+) \;=\; 1-\exp(-0.0115y^+) \tag{1.94}$$

is introduced in the eddy-viscosity law, which now reads

$$\nu_t \;=\; C_\nu \; f_\nu \; \frac{k^2}{\epsilon} \; . \tag{1.95}$$

The figure 1.7 shows a log-lin plot of $f_\nu(y^+)$. Out of the viscous sublayer, for $y^+ \gtrsim 4$, the profile of $f_\nu(y^+)$ is, naturally, rather similar to the ones of the functions c_ν and R_ν that we introduced in the ultimate part of the problem 1.1, see the equations (1.92) and (1.93), and the figures 1.6fh...

[10] The value of 100 quoted here is somehow arbitrary... some start at lower or larger y^+...

[11] The coefficient of y^+ in the exponential is taken from Chien (1982), but Hanjalić & Launder (2011) report a 10 times large value, we have no explanation for this discrepancy.

It is worthwhile to try to develop an understanding of all these issues concerning the ***boundary conditions near a solid wall***, and of the related ***'wall-damping problems'***. These problems are rather tough, and we will unfortunately not solve them completely in this module. As a first step towards a solution, we focus in the forthcoming section on the ***behaviour of the turbulent fields in the viscous sublayer***.

1.8 Study of the turbulent fields in the viscous sublayer

Problem 1.2 ***The turbulent fields in the viscous sublayer: asymptotic study and numerical validation on turbulent channel flow***

We focus on the ***viscous sublayer*** near a ***plane wall*** defined by $y = 0$, that has been already studied in channel flows in the parts 1 and 2 of the problem 1.1: see the figures 1.1b for a sketch of the geometry and of the flow considered, 1.2 for the confirmation in DNS, as far as the mean flow is concerned. The study is performed for a general turbulent flow, where the ensemble average that appears for instance in the first of the equations (1.97) depends in principle on x, z and t, but validations and illustrations are focussed on ***turbulent channel flow***, using the DNS database (1.1). There the ensemble averages do not depend on x and z, i.e., the turbulence is 'homogeneous' in these directions, and on t, i.e., the turbulence is 'stationary'.

1 With the help of the boundary conditions and of a conservation equation, show that the components of the ***fluctuating velocity*** **u** behave as follows, as $y \to 0$,

$$\begin{array}{rcccl} u_x & = & A(x,z,t)\ y & + & O(y^2)\ , \\ u_y & = & B(x,z,t)\ y^2 & + & O(y^3)\ , \\ u_z & = & C(x,z,t)\ y & + & O(y^2)\ . \end{array} \tag{1.96}$$

2.1 Deduce from this the behaviour of the ***velocity (co)variances*** as $y \to 0$,

$$\begin{array}{rcccl} \overline{u_x u_x} & = & \overline{A^2(x,z,t)}\ y^2 & + & O(y^3)\ , \\ \overline{u_y u_y} & = & & + & O(\quad)\ , \\ \overline{u_z u_z} & = & & + & O(\quad)\ , \\ \overline{u_x u_y} & = & & + & O(\quad)\ . \end{array} \tag{1.97}$$

Comment briefly, in connection to the figures 1.4abce.

2.2 In a turbulent channel flow at a given Reynolds number, because of the homogeneity in the directions x and z and in time, show that there exist constants A_0 , B_0 , C_0 and D_0 such that, as $y \to 0$, in wall units,

$$\overline{u_x^+ u_x^+} \sim A_0\ y^{+2}\ , \quad \overline{u_y^+ u_y^+} \sim B_0 \quad , \quad \overline{u_z^+ u_z^+} \sim C_0 \quad , \quad -\overline{u_x^+ u_y^+} \sim D_0 \quad . \tag{1.98}$$

Validate the formulas regarding $\overline{u_x^+ u_x^+}$, $\overline{u_y^+ u_y^+}$ and $\overline{u_x^+ u_y^+}$ by realizing, for $Re_\tau = 180$, the plots asked for in the figures 1.8abc. The constants A_0 etc. will always be computed by assuming that, in (1.98), the sign $\sim$ may be replaced by the sign $=$ at the first grid point off the wall. For this purpose, restart from the **Matlab** program that you used to solve the question 5.1 of the problem 1.1, working on the file `LM_Channel_180_vel_fluc_prof.dat`, and use the `loglog` command for log-log scales. To better compare the figures 1.8abc, always use the same ranges of values of y^+ and of the ordinates Var,

$$y^+ \in [0.1, 100]\ , \quad \text{Var} \in [10^{-7}, 10]\ . \tag{1.99}$$

Comment briefly.

3.1 Deduce from this the behaviour of the ***turbulent kinetic energy*** k as $y \to 0$,

$$k \;=\; \qquad\qquad + \; O(\quad) \; . \tag{1.100}$$

3.2 In a turbulent channel flow at a given Reynolds number, show that there exists a constant K_0 such that, as $y \to 0$, in wall units,

$$k^+ \;\sim\; K_0 \; y^{+2} \; . \tag{1.101}$$

Validate this formula by realizing the plot asked for in the figure 1.8d, still using the ranges (1.99), for $Re_\tau = 180$. For this value of Re_τ , precise the value of K_0 estimated from the first grid point off the wall,

$$K_0 \;=\; \qquad . \tag{1.102}$$

4.1 We now focus on the ***turbulent dissipation rate***

$$\varepsilon \;=\; \nu \; \overline{\varepsilon'} \qquad \text{with} \qquad \varepsilon' \;=\; (\partial_{x_j} u_i) \; (\partial_{x_j} u_i) \; . \tag{1.103}$$

Using the asymptotic expansions (1.96), show that for 7 couples (j,i) the square $(\partial_{x_j} u_i)^2$ is $O(y^2)$ as $y \to 0$, whereas for 2 couples (j,i) the square $(\partial_{x_j} u_i)^2$ has a finite limit as $y \to 0$. Deduce from this, in coherence with what the figure 1.5f displays, that ε has a finite limit as $y \to 0$, i.e., a ***value at the wall***, which is a simple function of ν and A, B, C introduced in (1.96),

$$\varepsilon_w \;:=\; \lim_{y \to 0} \varepsilon \;=\; \qquad\qquad . \tag{1.104}$$

4.2 Show that one has the relation, that may be viewed as a ***boundary condition*** for the ***turbulent dissipation rate*** ε,

$$\boxed{\varepsilon_w \;=\; \nu \; \partial_y^2 k \;=\; \lim_{y \to 0} \frac{2\nu k}{y^2}} \; , \tag{1.105}$$

i.e., in wall units,

$$\boxed{\varepsilon_w^+ \;=\; \partial_{y^+}^2 k^+ \;=\; \lim_{y^+ \to 0} \qquad} \; . \tag{1.106}$$

4.3 Check that, for the turbulent channel flow DNS at $Re_\tau = 180$, the expected value

$$\varepsilon_w^+ \;=\; \qquad = \qquad \; , \tag{1.107}$$

according to the equation (1.102), fits to the corresponding curve of the figure 1.5f and to the first data point of the corresponding column of the file `LM_Channel_0180_RSTE_k_prof.dat`.

a : **b :**

c : **d :**

e :

Fig. 1.8 : DIY ! Various turbulent fields extracted from the DNS at $Re_\tau = 180$ of the database (1.1), in dimensionless wall units, plotted vs y^+ in log-log scales. **a :** continuous curve: the ***variance*** $\overline{u_x^+ u_x^+}$; dashed line: the power law $A_0 y^{+2}$, with A_0 computed from $\overline{u_x^+ u_x^+} = A_0 y^{+2}$ at the first grid point off the wall. **b :** continuous curve: the ***variance*** $\overline{u_y^+ u_y^+}$; dashed line: the power law $B_0 y^{+4}$. **c :** continuous curve: the opposite of the ***covariance*** $-\overline{u_x^+ u_y^+}$; dashed line: the power law $D_0 y^{+3}$. **d :** continuous curve: the ***turbulent kinetic energy*** k^+; dashed line: the power law $K_0 y^{+2}$. **e :** continuous curve: the ***eddy viscosity*** ν^+; dashed line: the power law $D_0 y^{+3}$.

5.1 We now focus on turbulent flows for which the ***mean flow*** is ***unidirectional***,

$$\mathbf{V} \;=\; U(y)\;\mathbf{e}_x \;, \tag{1.108}$$

hence the ***eddy viscosity*** may be estimated like in the question 5.4 of the problem 1.1,

$$\nu_t \;=\; -\,\frac{\overline{u_x u_y}}{S} \quad \text{with the } \textbf{\textit{mean strain rate}} \quad S \;=\; \frac{dU}{dy}\;. \tag{1.109}$$

Using the mean-flow theory for the viscous sublayer[12], and the result of question 2.1, given by the last of the equations (1.97), establish the behaviour of the ***eddy viscosity*** as $y \to 0$,

$$\nu_t \;=\; -\qquad\qquad\qquad\qquad + \;O(\quad)\;. \tag{1.110}$$

5.2 Transform these relations to wall units, and to a turbulent channel flow case, using the notations of the equations (1.98), to show that, as $y \to 0$,

$$\nu^+ \;=\; -\,\frac{\overline{u_x^+ u_y^+}}{S^+} \;\sim\qquad\qquad\sim\qquad\qquad . \tag{1.111}$$

Validate this relation for the DNS at $Re_\tau = 180$ by realizing the plots asked for in the figure 1.8e.

5.3 Still for a turbulent channel flow case, compare this asymptotic power law to the one derived from the eddy-viscosity law of the $k-\varepsilon$ model, in connection with what display the figures 1.6fh.

Comments on the problem 1.2

The issue of the different asymptotic behaviours of the eddy viscosity of the DNS and the (ideal) $k-\varepsilon$ model in the viscous sublayer, evidenced in your answer to the question 5.3, is not very important, since the eddy viscosity plays no role there, as explained in your answer to the question 7.3 of the problem 1.1.

On the contrary, the fact that we established in question 4.2, equation (1.105), a physically sound boundary condition at a solid wall on the second turbulent field, ε, of the $k-\varepsilon$ model, is an important result of the problem 1.2. However, this condition is not convenient from a computational point of view, since it couples ε at the wall to the second derivative of k at the wall. Moreover, it can be shown that the equation (1.105) is totally incompatible with the high-Reynolds number ε - equation (1.45) of the $k-\varepsilon$ model... which is bad news for this model !..
If one considers, as the second turbulent field, instead of ε, the ***turbulent frequency***

$$\boxed{\Omega \;:=\; \frac{\varepsilon}{k}}\;, \tag{1.112}$$

the equation (1.105) shows that, near a solid wall, as the wall-normal coordinate $y \to 0$, one has the ***'boundary condition'***

$$\boxed{\Omega \;\sim\; \frac{2\nu}{y^2}}\;. \tag{1.113}$$

This may be enforced at the first grid points off the wall, placed in the viscous sublayer i.e. with $y^+ < 4$, replacing there the sign $\sim$ by the sign $=$. The fully explicit 'boundary condition' (1.113) is simpler than the boundary condition expressed by the equation (1.105). By this remark, we are led to the $k-\omega$ ***models***, which are interesting alternatives to the $k-\varepsilon$ model...

[12] See the part 1 of the problem 1.1 and the comments on parts 1 to 3 of this problem.

1.9 The $k-\omega$ models

1.9.1 The $k-\omega$ model of Wilcox (1988)

The idea of using the ***turbulent kinetic energy*** k and the ***turbulent frequency*** Ω as the turbulent fields that determine the eddy viscosity dates back to Kolmogorov (1942). In this seminal paper, he proposed the first two-equations eddy-viscosity model of the History of science, and this was a $k-\omega$ model. This model was revisited and improved by Wilcox (1988). We now give an overview of his model, using his notations, but in the case of incompressible fluids.
Wilcox (1988) chose to define the ***turbulent frequency*** or ***specific dissipation rate*** (see the discussion around the equation 1.118 below) as

$$\boxed{\omega \;:=\; \frac{\varepsilon}{\beta^* k} \;=\; \frac{\Omega}{\beta^*}} \tag{1.114}$$

with the ***eddy-viscosity coefficient***

$$\boxed{\beta^* \;=\; C_\nu \;=\; 0.09}\;, \tag{1.115}$$

such that the ***eddy-viscosity*** is given by the law equivalent to the one of the $k-\varepsilon$ model, equation (1.43),

$$\boxed{\nu_t \;=\; \frac{k}{\omega}}\;. \tag{1.116}$$

The ***modeled turbulent kinetic energy equation*** is similar to the one of the $k-\varepsilon$ model (1.44),

$$\boxed{\frac{Dk}{Dt} \;=\; \underbrace{\partial_{x_j}\Big((\nu+\sigma^*\nu_t)\partial_{x_j}k\Big)}_{D^\nu+D^t} \;+\; P_k \;-\; \beta^* k\omega}\;, \tag{1.117}$$

with σ^* a dimensionless number of order 1 (which equals in principle $1/\sigma_k$ of the $k-\varepsilon$ model), P_k the production term (1.32) or (1.33), and where the dissipation term ε has been replaced by its expression in terms of ω and k, see the equation (1.114). Importantly, in homogeneous turbulence, where in equation (1.117) the advection term in Dk/Dt vanishes, the diffusion and production terms also vanish, we get

$$\frac{\partial_t k}{k} \;=\; -\,\beta^*\omega\;. \tag{1.118}$$

For this reason Wilcox (1988) denote ω as the ***specific dissipation rate***.
Similarly to what Spalding and coworkers did in order to write and equation for ε, see our section 1.7.3, the ***modeled turbulent frequency equation*** is written by analogy to the equation for k, with viscous diffusion, turbulent diffusion, production and dissipation terms that have the correct dimension by multiplication by the factor ω/k for the last two terms, inserting also numbers to have new degrees of freedom for good modeling:

$$\boxed{\frac{D\omega}{Dt} \;=\; \underbrace{\partial_{x_j}\Big((\nu+\sigma\nu_t)\partial_{x_j}\omega\Big)}_{D^\nu_\omega+D^t_\omega} \;+\; \gamma\,\frac{\omega}{k}\,P_k \;-\; \beta\,\omega^2}\;. \tag{1.119}$$

On the basis of the study of representative flow cases, Wilcox (1988) proposed these values of the ***model coefficients***

$$\boxed{\beta^* \;=\; 0.09\;,\quad \sigma^* \;=\; \sigma \;=\; 0.5\;,\quad \gamma \;=\; 0.555\;,\quad \beta \;=\; 0.075}\;. \tag{1.120}$$

In particular, γ has been calculated by studying a ***log layer*** in a manner equivalent to what we did in the part 4 of the problem 1.1, see the exercise 1.1 below. From this analysis, ***wall laws*** for k and ω may be derived, that follow directly from the equations (1.73) and (1.114), i.e.,

$$k \;=\; \frac{u_\tau^2}{\sqrt{\beta^*}}\;,\quad \omega \;=\; \frac{u_\tau}{\sqrt{\beta^*}\;\kappa\;y}\;. \tag{1.121}$$

It may be checked that the exact 'near-wall boundary condition' that follows from the equations (1.113) and (1.114), namely

$$\omega \;\sim\; \frac{2\nu}{\beta^* y^2} \tag{1.122}$$

as the wall-normal coordinate $y \to 0$, is qualitatively compatible with the ω - equation (1.119). Indeed, an asymptotic analysis of the ω - equation (1.119) as $y \to 0$ shows that the dominant terms are the viscous diffusion and dissipation terms, that must balance. This balance leads to the ***'near-wall boundary condition'***

$$\omega \;\sim\; \frac{6\nu}{\beta y^2} \tag{1.123}$$

as $y \to 0$, which is quantitatively different from (1.122) but qualitatively similar, since the same power laws in $\nu^1\;y^{-2}$ are found.

Wilcox (1988) computed a unidirectional mean-flow

$$\mathbf{V} \;=\; U(y)\;\mathbf{e}_x\;, \tag{1.124}$$

using the near-wall boundary condition (1.123), together with the natural no-slip boundary conditions

$$U \;=\; k \;=\; 0 \tag{1.125}$$

at the wall ($y = 0$), and the log-layer laws (1.121) as $y \to +\infty$ (in practice, $y^+ = 500$). This gave as output a mean-flow $U(y)$ profile that approached, for $y^+ \gtrsim 20$, the log-law (1.61), with a reasonable value of the additive constant $C \simeq 5.1$. This success allowed him to declare that his model may be integrated down to the wall, i.e., through the viscous sublayer. This is a strong argument in favour of his model, which is considered therefore to be more relevant than the $k-\varepsilon$ model in near-wall regions.

1.9.2 Exercise

Exercise 1.1 ***Studies around the*** $k-\omega$ ***models***

We focus on $k-\omega$ RANS models. We consider a ***turbulent flow near a plane wall***, defined by $y=0$, such that, in the region of focus, the ***mean flow***

$$\mathbf{V} \;=\; U(y)\;\mathbf{e}_x\;, \tag{1.126}$$

the ***turbulent kinetic energy***

$$k \;=\; k(y)\;, \tag{1.127}$$

and the ***turbulent frequency***

$$\Omega \;=\; \Omega(y) \quad \text{or} \quad \omega \;=\; \omega(y) \tag{1.128}$$

depending on the normalization.
In part 1, we consider the ***near wall 'boundary condition'*** on Ω, and therefore focus on the ***viscous sublayer***.
In part 2, we study the $k-\omega$ ***model of Wilcox (1988)***, in the ***log layer***.

Part 1
Numerical study of the viscous sublayer on the turbulent channel flow DNS database

We consider the ***viscous sublayer*** as characterized in the parts 1 and 2 of the problem 1.1. We want to test on the same ***turbulent channel flow DNS database*** the ***'boundary condition'*** (1.113) for $\Omega=\varepsilon/k$.

1.1 Define Ω^+, Ω in ***wall units***, in terms of Ω. For this purpose, define a natural wall unit of time t_τ .
Check that Ω^+ is ε^+/k^+ with ε^+ and k^+ as defined in the lecture notes.

1.2 Establish the form of the 'boundary condition' (1.113) in wall units,

$$\boxed{\Omega^+ \;\sim\; \frac{2}{y^{+2}} \quad \text{as} \quad y^+ \to 0}\;. \tag{1.129}$$

1.3 Define a simple function $f(y^+)$ such that, if the 'boundary condition' (1.129) would hold exactly in the whole channel, then the ***premultiplied turbulent frequency***

$$f(y^+)\;\Omega^+(y^+) \tag{1.130}$$

would be equal to 1.

1.4 With **Matlab**, verify the relevance of the 'boundary condition' (1.129) on the DNS of the database at $Re_\tau=180$, by plotting $f(y^+)\;\Omega^+(y^+)$ vs y^+ in the interval $[0,10]$, and the ordinate in a relevant interval: realize the plot asked for in the figure 1.9a.
Comment. In particular, explain briefly why the first data point, at the lowest value of y^+, should be disregarded.

a : **b :**

Fig. 1.9 : DIY ! Study of the behaviour of the ***turbulent frequency*** $\Omega = \varepsilon/k$ in the ***viscous sublayer***, on the ***turbulent channel flow DNS*** of the database (1.1) at $Re_\tau = 180$. **a :** The premultiplied turbulent frequency $\frac{1}{2} y^{+2}\ \Omega^+$ (thick continuous curve and data points) and the limit 1 (thin horizontal dashed line). **b :** The error (1.131) (thick line and data points), the limit 0.07 (thin horizontal dashed line), and the integers y_1^+ and $y_1^+ + 1$ (thin vertical lines), see (1.132).

1.5 Expand your **Matlab** script to create the figure 1.9b that displays the error

$$\mathrm{err}(y^+) \ = \ |f(y^+)\ \Omega^+(y^+) - 1| \tag{1.131}$$

vs y^+ in the interval $[0, 10]$. Determine the largest integer y_1^+ such that,

$$\forall y^+ \in [0, y_1^+] \ , \quad \mathrm{err}(y^+) \ < \ 0.07 \ . \tag{1.132}$$

Comment.

Part 2

Analytical study of the log layer in the framework of the model of Wilcox (1988)

We consider the ***log layer*** as characterized in the part 1 of the problem 1.1. In particular, we assume that the mean-flow derivative and the dynamic eddy viscosity

$$\frac{dU}{dy} \ = \ \frac{u_\tau}{\kappa y} \ , \quad \eta_t \ = \ \rho u_\tau \kappa y \ , \tag{1.133}$$

with u_τ the ***friction velocity*** and κ the ***von Karman constant***.

2.1 Explicit the components of the ***mean gradient of velocity*** and ***mean rate-of-strain tensors***, depending on the couple of cartesian indices considered. In the framework of the $k-\omega$ model, calculate the ***(turbulence) production term*** P_k .

2.2 Assuming that the diffusion terms in the $k-\omega$ model equation for k are negligible in front of the ***production*** and ***dissipation terms***, establish a relation between k and ω. Comment briefly.

2.3 By using also the $k-\omega$ ***eddy-viscosity law***, calculate k and ω in the log layer. Comment.

2.4 Show that the model equation for ω is valid in the log layer, provided that a reasonable approximation is made, and there is a relation between the von Karman constant κ and some coefficients of the model $k-\omega$,

$$\kappa^2 \ = \qquad\qquad . \tag{1.134}$$

2.5 Verify that the coefficients of the model $k-\omega$ (1.120) correspond to a reasonable value of κ . Comment briefly and from an 'historical' point of view.

1.9.3 Opening: the $k-\omega$ Shear-Stress Transport model of Menter (1994)

With the idea that the $k-\omega$ model is more relevant in near-wall regions, Menter (1994) proposed a model that 'blends' the $k-\omega$ and $k-\varepsilon$ models, the $k-\omega$ model being recovered in the near-wall region, the $k-\varepsilon$ model in the freestream region. With a further, clever modification of the eddy-viscosity law, this gave the 'famous' $k-\omega$ ***Shear-Stress Transport*** **(SST)** ***model***, which is still widely used nowadays in CFD studies...

1.10 Openings

Many RANS models exist, that have not been described nor even advocated in the core of this chapter. An illustration of this profusion of models is provided by a glimpse at the ***'Turbulence Modeling Resource'*** web site of NASA Langley Research Center,

`https://turbmodels.larc.nasa.gov` . (1.135)

This page offers a list of 18 'popular' RANS models, which are all (follow the links !) described in some details, and then tested on some benchmark cases. From a more theoretical point of view, a weak point with the RANS approach is that most (if not all) model equations have been constructed from physical and dimensional arguments, which is not very systematic. I am happy to state that I contributed recently to a systematic derivation of a RANS eddy-viscosity equation, in a restricted class of flow, however, see Plaut & Heinz (2021)... which has to do with the problem 1.3 !.. This small step forward does not solve everything: it should be clear that other topics could be presented within the RANS framework, not speaking of other approaches that are currently spreading from the community of the researchers to the one of the engineers: the ***Large Eddy Simulations*** **(LES)** and the ***Hybrid RANS - LES methods***. These last methods might be presented as the best ones nowadays for the CFD of complex turbulent flows, with new ideas to better fix the models, at least this is what claims Heinz (2020)...

1.11 Problem: turbulent channel flow at high Reynolds number

We offer a problem that makes use of many concepts introduced in this chapter, but with a different point of view: the aim is to set up an asymptotic theory of turbulent channel flow in the near wall region and in the limit $Re_\tau \to +\infty$.

Problem 1.3 ***Analytic models of turbulent channel flow in the near wall region at high Reynolds numbers***

Inspired by Heinz (2018, 2019), we want to ***model analytically turbulent channel flow*** in the limit $Re_\tau \to \infty$ and in the ***near wall region***. We use exactly the same notations as in the problems 1.1 and 1.2, except for the (main i.e. non-vanishing) ***Reynolds shear stress*** which is denoted τ instead of τ_{xy} , and for the ***production term*** in the ***turbulent kinetic energy equation***, which is denoted P instead of P_k . We recall that in wall units

$$\tau^+ \;=\; \frac{\tau}{\tau_w} \;=\; -\,\overline{u_x^+ u_y^+}\;. \tag{1.136}$$

Part 1 ***Model of the Reynolds shear stress***

We want to show that the analytic expression

$$\tau_m^+ \;=\; \left(\frac{(y^+/a)^{b/c}}{1+(y^+/a)^{b/c}} \right)^c \tag{1.137}$$

is a good model for τ^+ in the near wall region, provided that the parameters $a,\ b$ and c are set correctly in $\mathbb{R}^{+*}$.

1.1 As $y^+ \to 0$, show that

$$\tau_m^+ \;\sim\; D_0\ y^{+n} \tag{1.138}$$

with D_0 a coefficient and n an exponent that depend simply on the parameters $(a,\ b)$,

$$D_0 \;=\; \qquad ,\quad n \;=\; \qquad . \tag{1.139}$$

Explain how a may be calculated from D_0 and b.

1.2 Using one result of the problem 1.2, fix the (integer) value of b,

$$b \;=\; \qquad . \tag{1.140}$$

1.3 To fix the value of a, we use the DNS of the database (1.1) at the highest Re_τ available, $Re_\tau = 5200$. Using **Matlab** and the file `LM_Channel_5200_vel_fluc_prof.dat`, extract, from the knowledge of τ^+ at the first grid point off the wall, an estimate of the numerical value of D_0 such that, as $y^+ \to 0$,

$$\tau^+ \;\sim\; D_0\ y^{+n} \tag{1.141}$$

with n already fixed according to the previous questions. Fix from this the value of a; give the values of $y^+,\ D_0$ and a thus found, with 3 digits[13]:

$$y^+ \;=\; \qquad ,\ D_0 \;=\; \qquad ,\ a \;=\; \qquad . \tag{1.142}$$

Using other files of the DNS database, realize the figure 1.10a. For all the 4 figures 1.10a to 1.12 that you have to realize, set the interval of the abscissa as follows,

$$y^+ \;\in\; [0.1,\ 5200]\ .$$

In all figures, mark one or a few relevant (figure - dependent !) values of y^+ with vertical lines, and if necessary a relevant value of the ordinate with an horizontal line.
Comment the figure 1.10a.

1.4 To fix the value of c, explain on the basis of the figure 1.10a why enforcing $\tau^+ = \tau_m^+$ at a value of y^+ close to 14 is relevant. By inspecting with **Matlab** the tabulated values of y^+ for the DNS at $Re_\tau = 5200$, extract the index i such that the i^{th} value of y^+ is the one closest to 14. Give $i,\ y^+$ and τ^+ thus found, with 3 digits:

$$i \;=\; \qquad ,\quad y^+ \;=\; \qquad ,\quad \tau^+ \;=\; \qquad . \tag{1.143}$$

[13] To force **Matlab** to print more digits, use `format long` at the beginning of your script.

a : **b :**

Fig. 1.10 : DIY ! **a :** The ***Reynolds shear stress*** τ^+ vs y^+, in log-log scales, for the ***channel flow DNS*** at $Re_\tau = 180$ in black, 1000 in blue, 5200 in red, with continuous lines, and the power law (1.138) with a black dashed line. The vertical lines mark $y^+ =$, 14 and , the horizontal line marks $\tau^+ =$. **b :** Same as a, except that the black dashed line now shows the complete analytic model τ_m^+ , and that the vertical lines mark $y^+ =$ and .

1.5 Since the equation $\tau^+ = \tau_m^+$ implies c in a complex nonlinear manner, create a **Matlab** script `taum.m` which codes the function τ_m^+ as follows,

```
function res= taum(yplus,a,b,c)
 res= ( ((yplus/a)^(b/c)) / ( 1 + (yplus/a)^(b/c)) )^c;
```

and a script `differ.m` which codes the function $differ = \tau_m^+ - \tau^+$ as follows,

```
function res= differ(yplus,a,b,c,taudns)
  res = taum(yplus,a,b,c) - taudns;
```

Check that, for the values of y^+ and τ^+ determined in question 2.1, and with the previously fixed values of a and b, the function *differ* changes of sign in the interval $c \in [1, 2]$. Fix therefore the value of c by computing the root of the equation $\tau^+ = \tau_m^+$ that sits in this interval as follows,

```
c= fzero(@(x) differ(yplus,a,b,x,taudns), [1 2])
```

Give this value with 2 digits only, since it turns out that the model is not very sensitive to the value of c,

$$c = \quad . \tag{1.144}$$

Realize the figure 1.10b; to plot the profile of the model, use the command `arrayfun` to apply the function `taum` to the tabulated values of y^+, according to something like

```
tau3m= arrayfun(@(x) taum(x,a,b,c),yp3)
```

Comment this figure.

Fig. 1.11 : DIY ! The ***mean strain rate*** S^+ vs y^+, in log-log scales, for the ***channel flow DNS*** at $Re_\tau = 180$ in black, 1000 in blue, 5200 in red, with continuous lines, and the model $S_m^+ = 1 - \tau_m^+$ with a black dashed line. The vertical line marks $y^+ =$. **For bonus - optional:** the black dotted line shows the alternate log layer model $S_{\ell\ell}^+$.

Part 2 *Model of the mean strain rate*

2.1 Show that the dimensionless version, in wall units, of the RANS momentum equation (1.49) links the ***mean strain rate***

$$S^+ = \frac{dU^+}{dy^+} \tag{1.145}$$

with τ^+ and y/δ.

2.2 Explain why it is reasonable to assume that, at high Reynolds number, at fixed y^+ in the near wall region,

$$S^+ = 1 - \tau^+ . \tag{1.146}$$

2.3 From this and part 1 an analytical model for S^+ is given by

$$S_m^+ = 1 - \tau_m^+ . \tag{1.147}$$

Using some files `LM_Channel_*_mean_prof.dat` of the DNS database, realize with **Matlab** the figure 1.11, and comment it.

2.4 For bonus - optional ! Calculate the mean strain rate $S_{\ell\ell}^+$ that should be obtained in the log layer, assuming that the mean flow would be exactly given by the equation (1.61),

$$U^+ = U_{\ell\ell}^+ = \frac{1}{\kappa} \ln y^+ + C .$$

Add a line showing $S_{\ell\ell}^+$ on the graph of the figure 1.11, using the value of κ measured in the problem 1.1, and comment briefly the relevance of this alternate model for S^+.

Fig. 1.12 : DIY ! The ***turbulent kinetic energy production term*** P^+ vs y^+, in log-linear scales, for the ***channel flow DNS*** at $Re_\tau = 180$ in black, 1000 in blue, 5200 in red, with continuous lines, and the model P_m^+ with a black dashed line. The vertical lines mark $y^+ = y_2^+$ and 30, the horizontal line marks $P^+ = 1/4$.

Part 3 ***Model of the turbulent kinetic energy production term***

3.1 We recall that the ***production term*** in the ***turbulent kinetic energy equation*** reads, in wall units,

$$P^+ \;=\; \tau^+ \, S^+ \; . \tag{1.148}$$

Show that, at high Reynolds number, in the near wall region, P^+ is a simple polynomial function of degree 2 of S^+,

$$P^+ \;=\; \qquad\qquad\qquad . \tag{1.149}$$

Deduce from this, since S^+ varies from 1 at the wall to 0 at the centerplane, that $\max P^+ = 1/4$ for a well defined value $S_2^+ \in \mathbb{Q}$ of S^+,

$$S_2^+ \;=\; \qquad . \tag{1.150}$$

3.2 Using the model S_m^+ of S^+ defined in part 2, calculate analytically, as a function of the parameters (a, b, c), the value y_2^+ of y^+ where $S_m^+ = S_2^+$,

$$y_2^+ \;=\; \qquad\qquad\qquad . \tag{1.151}$$

3.3 With **Matlab**, using the values of (a, b, c) determined in part 1, give the numerical value of y_2^+, with 3 digits, and comment it briefly,

$$y_2^+ \;=\; \qquad . \tag{1.152}$$

3.4 Using the analytical models τ_m^+ of τ^+ and S_m^+ of S^+ of parts 1 and 2 defines an analytical model of P^+,

$$P_m^+ \;=\; \tau_m^+ \, S_m^+ \; . \tag{1.153}$$

With **Matlab**, using some files `LM_Channel_*_RSTE_k_prof.dat` of the DNS database, realize the figure 1.12 and comment it.

Chapter 2

Wind energy and turbulence

This chapter corresponds to the sessions 4 *to* 6.

2.1 Wind energy: power performance theory

This section presents the concept of ***power performance for wind turbines***, starting from momentum theory to power curves that leads to the Betz limit. This section deals only with horizontal-axis three-bladed electrical wind turbines. There is no major limitation to its extension to other designs of wind power systems.

2.1.1 Momentum theory for wind turbines

A basic understanding of fluid mechanics will be applied to wind turbines, with the so-called 'momentum theory'. This theoretical approach sets ground for the further power curve analysis. The complexity of turbulence is first set aside, so as to understand the fundamental behavior of a wind turbine in a uniform flow at steady-state. More complex atmospheric effects will be addressed later on.

As a wind turbine converts the power from wind into available electrical power, one can assume the following relation

$$\boxed{P(u) \;=\; c_p(u)\; P_{wind}(u)}\,, \tag{2.1}$$

where $P_{wind}(u)$ is the power contained in the wind passing with speed u through the wind turbine, and $P(u)$ is the ***electrical power extracted***. The amount of power converted by the wind turbine is given by the ***power coefficient*** $c_p(u)$, which represents the efficiency of the machine. As the input $P_{wind}(u)$ cannot be controlled, improving power performance means increasing the power coefficient $c_p(u)$. The power contained in a laminar incompressible flow of mass m and density ρ moving along the $x-axis$ with constant speed u through a vertical plane of area A is

$$P_{wind}(u) \;=\; \frac{dE_{kin,wind}}{dt} \;=\; \frac{d}{dt}\left(\frac{1}{2}mu^2\right) \;=\; \frac{1}{2}\frac{dm}{dt}u^2 \;=\; \frac{1}{2}\frac{d(\rho V)}{dt}u^2 \;=\; \frac{1}{2}\rho\frac{d(Ax)}{dt}u^2 = \frac{1}{2}\rho A u^3\,, \tag{2.2}$$

as can be also deduced from the 'Euler formula' (see e.g. Plaut 2021*a*).

Let us consider a mass of air moving towards a wind turbine, which can be represented by an 'actuator disc' of diameter D. An actuator disc is an infinitely thin disc through which the air can flow without resistance, as proposed by Froude and Rankine's momentum theory Rankine (1865).

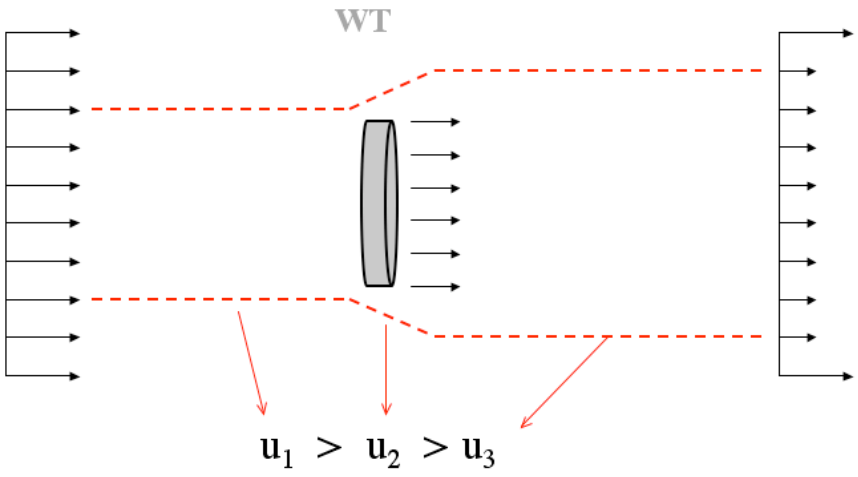

Fig. 2.1 : Idealized flow situation around a wind turbine (WT) according to Twele & Gasch (2005). The dashed lines show the limit of the 'streamtube' considered in the momentum theory, i.e., the stream of air particles that interact with the wind turbine. The wind speeds before, at and after the wind turbine are respectively u_1, u_2 and u_3.

When crossing the wind turbine, the wind is affected as parts of its energy is extracted. This extraction of kinetic energy results in a drop in the wind speed from upstream to downstream. The velocity far before the wind turbine (upstream), at the wind turbine and far behind (downstream) are labelled respectively u_1, u_2 and u_3. An illustration is given in figure 2.1. Mass conservation requires that the flow-rate $\dot{m} = A_i \rho u_i$ be conserved and

$$A_1 \rho u_1 \;=\; A_2 \rho u_2 \;=\; A_3 \rho u_3 \,, \tag{2.3}$$

where A_i are the respective areas perpendicular to the flow. A_2 is the area swept by the rotor blades $A_2 = A = \pi D^2/4$. As a consequence of the wind speed slowing down, i.e. $u_3 < u_2 < u_1$, the area of the stream-tube has to expand, and $A_3 > A_2 > A_1$. This can be observed in figure 2.1. Also, the energy extracted by the wind turbine can be determined by the difference of kinetic energy upstream and downstream of the wind turbine

$$E_{ex} \;=\; \frac{1}{2} m (u_1^2 - u_3^2) \,, \tag{2.4}$$

resulting in a power extraction

$$P_{ex} \;=\; \frac{d}{dt} E_{ex} \;=\; \frac{1}{2} \dot{m} (u_1^2 - u_3^2) \,. \tag{2.5}$$

The wind turbine continuously takes energy out of the wind flow, which reduces its velocity. However, the flow needs to *escape* the wind turbine downstream with a speed $u_3 > 0$. If all the power content of the wind would be extracted, the wind speed downstream would then become zero. As a consequence, the air would *accumulate* downstream and block newer air from flowing through the wind turbine, so that no more power could be extracted. This means that the wind flow must keep some energy to escape, which naturally sets a limit for the efficiency of any wind power system. The power coefficient $c_p(u)$ must be inferior to 1. An optimal ratio of wind speeds $\mu = u_3/u_1$ can be found that allows for the highest energy extraction.

2.1.2 Power performance - Betz limit

In the plane of the rotor blades, the flow-rate is given by

$$\dot{m} \;=\; \rho A u_2\,, \tag{2.6}$$

where the velocity u_2 can be determined following Froude-Rankine by calculating the thrust T given by

$$T \;=\; \dot{m}(u_1 - u_3)\,. \tag{2.7}$$

The corresponding power, based on the thrust P_{thrust}, is given by

$$P_{thrust} \;=\; \dot{m}(u_1 - u_3)u_2\,. \tag{2.8}$$

This power is equal to the extracted power P_{ex} from equation (2.5)

$$\dot{m}(u_1 - u_3)u_2 \;=\; \frac{1}{2}\dot{m}(u_1^2 - u_3^2)\,, \tag{2.9}$$

which can be solved for the velocity in the rotor plane leading to

$$u_2 \;=\; \frac{u_1 + u_3}{2}\,. \tag{2.10}$$

Inserting equations (2.10) and (2.6) into equation (2.5) yields

$$P(\mu) \;=\; \frac{1}{2}\rho A u_1^3\,\frac{1}{2}(1 + \mu - \mu^2 - \mu^3) \;=\; P_{wind}(u_1)\; c_p(\mu) \tag{2.11}$$

where $\mu = u_3/u_1$ is the wind speed reduction factor. The theoretical definition of the power coefficient is then

$$c_p(\mu) = \frac{1}{2}(1 + \mu - \mu^2 - \mu^3)\,, \tag{2.12}$$

as shown in figure 2.2. The optimal power performance is obtained for a ratio μ such that the derivative of $c_p(\mu)$ with respect to μ is zero

$$\frac{d}{d\mu}c_p(\mu) = \left(-\frac{1}{2}\right) \times \left(3\mu^2 + 2\mu - 1\right) = 0\,. \tag{2.13}$$

This leads to

$$\mu_{max} \;=\; \frac{1}{3} \quad\Longleftrightarrow\quad c_p(\mu_{max}) \;=\; \frac{16}{27} \;\simeq\; 0.593\,, \tag{2.14}$$

as shown in figure 2.2.

This limit is called the ***Betz limit***, see Betz (1927). In other words, a wind turbine can extract at most 59.3% of the power contained in the wind. This can be obtained when the wind speed downstream is one-third of the wind speed upstream.

A widely used representation of power performance is given by the relation of c_p to the ***tip speed ratio***

$$\lambda \;=\; \frac{\omega R}{u_1}\,, \tag{2.15}$$

where ω and R are the angular frequency and radius of the rotor. λ is the ratio of the rotational speed at the tip of the blades to the upstream wind speed. If the rotational speed increases, more air is 'slowed down' by the turbine, i.e., the wind speed reduction factor μ decreases: λ is directly related to μ. The dimensionless $c_p - \lambda$ curve will be introduced in the next section.

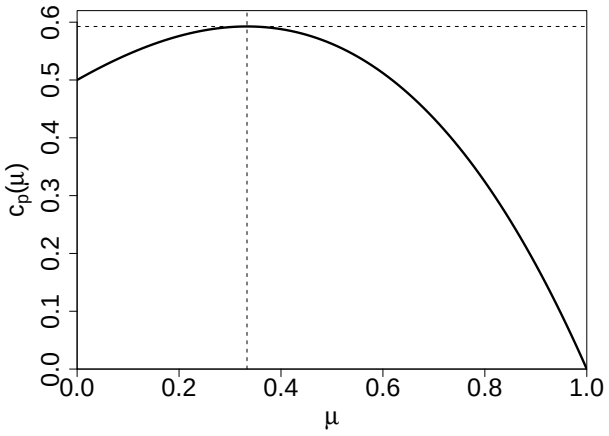

Fig. 2.2 : Power coefficient c_p as a function of the wind speed ratio $\mu = u_3/u_1$.

Betz' momentum theory only considers the mechanical transfer of energy from the wind to the rotor blades. The next step of the conversion from mechanical to electrical energy has not been taken into account, as well as all energy losses. The more complex design of wind turbines causes lower values of c_p, as discussed in section 2.1.4. The power coefficients of modern commercial wind turbines reach values of order 0.5. Also, criticism of Betz theory is given in Rauh & Seelert (1984); Rauh (2008), leading to a less well defined upper limit of c_p.

2.1.3 Energy conversion - Pressure drop

Wind turbines convert kinetic energy into electrical energy at the rotor plane. Therefore the question arises, which physical quantity is affected by this conversion process. Using the notation from figure 2.3 we can use Bernoulli equation to calculate the energy within the volume that approaches the wind turbine and which is left once the volume passed the turbine.

$$E_{in} = V_{in}\left(p_1 + \frac{1}{2}\,\rho\,u_1^2\right) = V_{in}\left(p_{-2} + \frac{1}{2}\,\rho\,u_{-2}^2\right) \tag{2.16}$$

$$E_{out} = V_{out}\left(p_3 + \frac{1}{2}\,\rho\,u_3^2\right) = V_{out}\left(p_{+2} + \frac{1}{2}\rho u_{+2}^2\right) \tag{2.17}$$

Assuming

$$\begin{aligned} V_{in} &= V_{out} \\ p_1 &= p_3 \\ u_{-2} &= u_{+2} \end{aligned}$$

equations (2.16) and (2.17) read

$$p_1 + \frac{1}{2}\,\rho\,u_1^2 = p_{-2} + \frac{1}{2}\,\rho\,u_{-2}^2 \tag{2.18}$$

$$p_1 + \frac{1}{2}\,\rho\,u_3^2 = p_{+2} + \frac{1}{2}\,\rho\,u_{-2}^2\,. \tag{2.19}$$

Solving equation (2.19) for $\frac{1}{2}\,\rho\,u_{-2}^2$ we get

$$\frac{1}{2}\,\rho\,u_{-2}^2 = p_1 + \frac{1}{2}\,\rho\,u_3^2 - p_{+2}$$

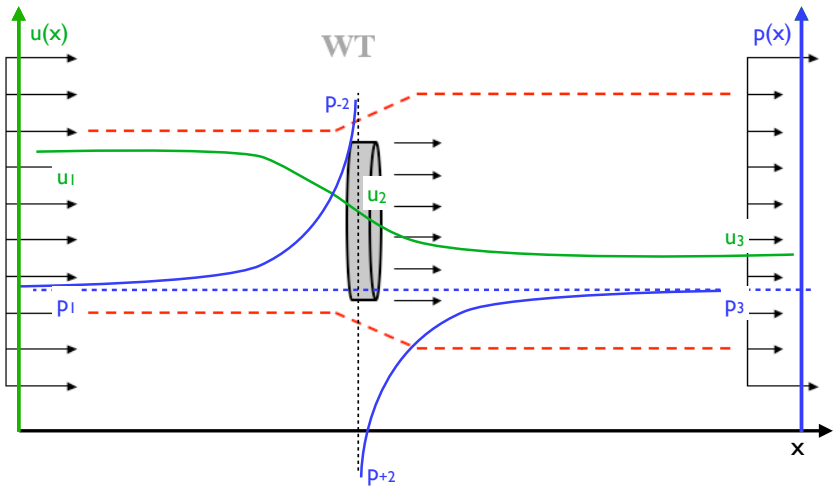

Fig. 2.3 : Idealized 'streamtube' with velocities u_1, $u_2 = \frac{2}{3}u_1$ and $u_3 = \frac{1}{3}u_1$ according to Betz limit. The green line indicates one possible behaviour of the velocity through the wind turbine. The blue curve sketches the behaviour of the static pressure, where p_1 denotes the pressure far in front of the turbine and p_3 the pressure far behind the turbine. p_{-2} denotes the pressure right before the rotor plane and p_{+2} the pressure right behind the rotor plane.

and substituting in equation (2.18) leads to

$$p_1 + \frac{1}{2}\,\rho\,u_1^2 \;=\; p_{-2} + p_1 + \frac{1}{2}\,\rho\,u_3^2 - p_{+2}\;. \tag{2.20}$$

Solving equation (2.20) for p_{+2}, which is the static pressure right behind the rotor plane reveals that the conversion of kinetic into electrical energy results in a pressure drop,

$$p_{+2} \;=\; p_{-2} - \frac{1}{2}\,\rho\left(u_1^2 - u_3^2\right)\;. \tag{2.21}$$

2.1.4 Limitations of Betz theory - Energy losses

Although it is based on a simplified approach, the Betz limit is a widely used and accepted value. But more realistic considerations indicate that real wind turbine designs have even lower efficiency due to additional limitations. In this section, the three main limitations to reach the optimal value of $c_p = 16/27$ are introduced. This section only aims to give a first idea. For a more detailed understanding of the mathematical equations presented here, the reader is kindly referred to the literature.

Bouncing losses

Betz' consideration does not take into account that there is not only a reduction of wind speed downstream, but also an additional angular momentum that is transferred to the air flow, as shown in figure 2.4. This effect follows Newton's third law, as a reaction to the rotational motion of the rotor. For slow rotating wind turbines (λ small), these losses are much more severe than for fast rotating machines. For $\lambda \approx 1$ an optimum value of c_p of only 0.42 can be reached instead of the Betz optimum of 0.59. c_p approaches the Betz optimum with increasing tip speed ratio.

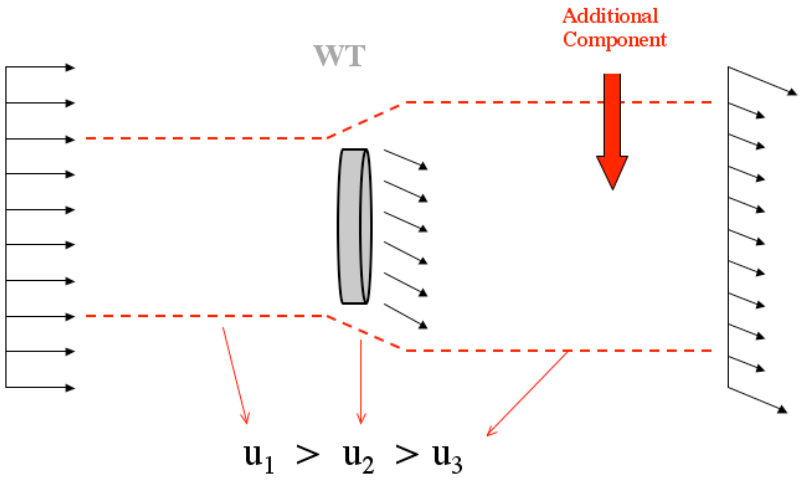

Fig. 2.4 : Flow around a wind turbine. After passing the turbine the velocity field has a rotational component due to the rotating rotor blades.

Profile losses

Another important source of energy loss is the quality of the airfoil profile for which the realistic or ideal cases, including drag force or not, can be considered.
The efficiency η can now be defined as the ratio between equation realistic and ideal situation. This leads to the efficiency given by

$$\eta \;=\; 1-\xi\,. \tag{2.22}$$

The profile losses ξ_{prof} follow the relation

$$\xi_{prof} \;\propto\; r\lambda\,. \tag{2.23}$$

In contrast to the bouncing losses, the profile losses mainly affect fast rotating machines. For higher tip speed ratios, the lift to drag ratio C_l/C_d must be optimized. Furthermore the losses increase with the radius, such that the manufacturing quality of the blade tips is of primary importance for power performance.

Tip losses

A good quality of the tips especially means that they should be as narrow as possible because this corresponds to an (ideal) airfoil with length infinity, $R/c \to \infty$ with R the distance to the rotation axis and c the chord. For real blades there is always a flow around the end of the blade (forming an eddy that is advected by the flow) from the high pressure area to the low pressure area. This is partly levelling the pressure difference and consequently the lift force. The tip losses obey approximately the following relation

$$\xi_{tip} \;\propto\; \frac{1}{z\lambda}. \tag{2.24}$$

Different to the profile losses an increasing tip speed ratio decreases the tip losses, as well as an increased number of blades.

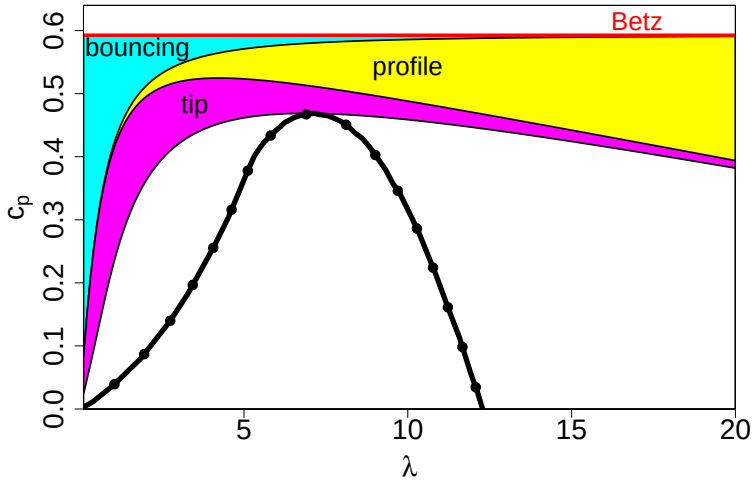

Fig. 2.5 : Typical $c_p - \lambda$ curve (black dotted line). The influence on the efficiency of Betz limit, the bouncing, profile and tip losses is illustrated. These results were obtained with a lift to drag ratio $C_L/C_D = 60$ and $z = 3$ blades. Maximum performance $c_p \simeq 0.475$ is obtained at $\lambda \simeq 7$.

Impact on power performance

The figure 2.5 shows an overview of the different kinds of losses and their influence on the value of c_p. One can see that bouncing losses cause the largest reduction in the power coefficient for small values of λ, similar to the finite number of blades. This is the opposite for the profile losses. Three-bladed wind turbines can reach optimal c_p values of order 0.50 for typical values of $\lambda \approx 6-8$, which naturally sets the strategy for optimal power performance in terms of rotational frequency ω.

2.1.5 Power curve

Along with the $c_p - \lambda$ curve, a standard representation of a wind turbine power performance is given by a so-called ***power curve***. The power curve gives the relation between the simultaneous wind speed u and power output P. Following usual practice, the ***wind speed*** u will refer to the upstream horizontal wind speed u_1 from now on, such that $u = u_1$. Also, the ***net electrical power output*** P that the wind turbine actually delivers to the grid is considered, integrating all possible losses. The two quantities u and P will follow these specifications until the end of the chapter. Following equation (2.11), the theoretical power curve reads

$$\boxed{P(u) \;=\; c_p(u)\; P_{wind}(u) \;=\; c_p(u)\; \frac{1}{2}\rho A u^3}\,. \tag{2.25}$$

In most of the modern wind turbine designs, the ***regulation of the power output*** is performed through changes both in the ***rotational frequency of the generator*** and in the ***pitch angle of each blade***. (Note other wind turbine designs involve fixed rotational frequency, so-called fixed-speed wind turbines, or fixed pitch angle, so- called fixed-pitch wind turbines). A more detailed description on control strategies is given in Bianchi *et al.* (2006). The rotational frequency of the generator is physically linked to the wind speed, such that it cannot be changed freely. However, the pitch angle of the blades can be controlled at will, and almost independently of the wind speed, to reach the chosen control strategy, and hence represents the central mean of control for the operation. Pitching plainly consists of a rotation of the blades by a pitch angle θ in the plane of their cross-section.

The power production is then controlled by changing the lift forces on the rotor blades (Burton *et al.* 2001; Bianchi *et al.* 2006). The power production can be reduced or stopped by pitching the blades towards stall[1]. In modern wind turbines, this is achieved by a so-called active pitch control. The power coefficient c_p depends strongly on this pitch angle θ and on the tip speed ratio λ, i.e. $c_p = c_p(\lambda(u), \theta)$. As λ can typically not be controlled, c_p is optimized via θ to a desired power production. In particular for high wind speeds, c_p is lowered to protect the wind turbine machinery and prevent from overshoots in the power production.

This pitch regulation is commanded by the controller of the wind turbine, which constitutes of several composite mechanical-electrical components that operate actively for the optimum power performance[2]. For the common pitch-controlled wind turbines, the control strategy gives four distinct modes of operation:

- for $u \leq u_{cut-in}$, here u_{cut-in} represents the minimum wind speed such that the wind turbine can extract power, typically in the order of $3-4$ m/s. In this range the power contained in the wind is not sufficient to maintain the wind turbine into motion, and no power is produced;

- in partial load $u_{cut-in} \leq u \leq u_{rated}$, here u_{rated} denoted the rated wind speed at which the wind turbine extracts the rated, maximum allowed power P_{rated}. u_{rated} is typically in the order of $12 - 15$ m/s. In this range the wind turbine works at its maximum power performance, i.e. c_p is maximized, and the pitch angle θ is normally maintained constant;

- in full load $u_{rated} \leq u \leq u_{cut-out}$, here $u_{cut-out}$ represents the maximum wind speed at which the wind turbine can safely extract power, typically in the order of $25 - 35$ m/s. In this range the wind turbine power output is limited to the rated power P_{rated}. In this mode of operation, the pitch angle θ is adjusted in real-time to maintain $P \approx P_{rated}$;

- for $u > u_{cut-out}$ the pitch angle θ is maximized to the feathered position so as to eliminate the lift forces on the blades. A braking device can be used in addition to block the rotation for safety reasons. As a consequence, the power production is stopped.

An illustration of the theoretical strategy for $c_p(u)$ and $P(u)$ is given in figure 2.6.

It is important to precise that this theoretical estimation is valid for a ***laminar flow***, which never occurs in real situations. The more complex atmospheric winds call for more complex descriptions of power performance. Following the path of turbulence research, statistical models are introduced in Appendix A to deal with this complexity.

2.2 Rotor blade: blade element momentum theory (BEM)

In this section we use the momentum theory for elements of the rotor blades to derive the ***design*** or layout of a blade.

[1] Stall effects are obtained when the angle of attack of an airfoil exceeds a critical value, resulting in a sudden reduction in the lift force generated.

[2] Additional considerations such as mechanical loads or power stability are usually taken into account as well Bianchi *et al.* (2006), but reach out of the scope of this chapter.

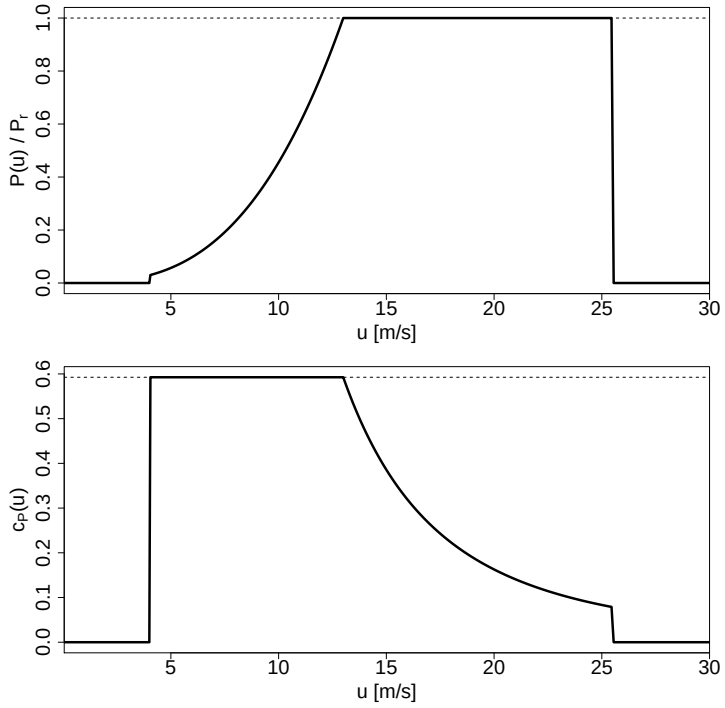

Fig. 2.6 : (a) Theoretical power curve $P(u)$; (b) Theoretical power coefficient $c_p(u)$ for a pitch-controlled wind turbine with $u_{cut-in} = 4$ m/s, $u_{rated} = 13$ m/s and $u_{cut-out} = 25$ m/s.

Modern wind turbines rotate due to the lift forces F_l acting on the airfoils. For an airfoil the effective area can be expressed in terms of the ***depth***, also called ***chord***, c, and the ***span*** of the wing b, similar to the rotor radius R. Therefore the ***drag force*** F_d and the ***lift force*** F_l read

$$F_d \;=\; C_d(\alpha)\frac{1}{2}\rho\, u^2\;(c\;b)\;, \quad F_l \;=\; C_l(\alpha)\frac{1}{2}\rho\, u^2\;(c\;b)\,, \tag{2.26}$$

where α is the ***angle of attack***, as displayed in figure 2.7. The lift-to-drag ratio F_l/F_d relates to the quality of the airfoil, which should be maximized.

Incident velocity

In figure 2.7, the velocity vector u_{res} gives the wind velocity in the frame of reference of the airfoil. The wind velocity at the rotor is $\frac{2}{3}u_1$ in the frame of the ground, where u_1 is the free wind velocity upstream in front of the turbine. Additionally, the rotational motion must be considered for the motion of the wind with respect to the blades. The velocity of the rotational motion at a radial position r is $u_{rot} = \omega r$, such that

$$u_{res}^2(r) = (2u_1/3)^2 + (\omega r)^2 \tag{2.27}$$

gives the effective or resulting velocity u_{res} of the air in the reference frame of the blade, see figure 2.7.

Forces on blade section dr

Blade element momentum theory is commonly used to estimate the total force acting on the rotor by summing up the local force on each infinitesimal blade element of size dr. The total force is divided into its rotational component F_{rot} and its axial component F_a, which causes the

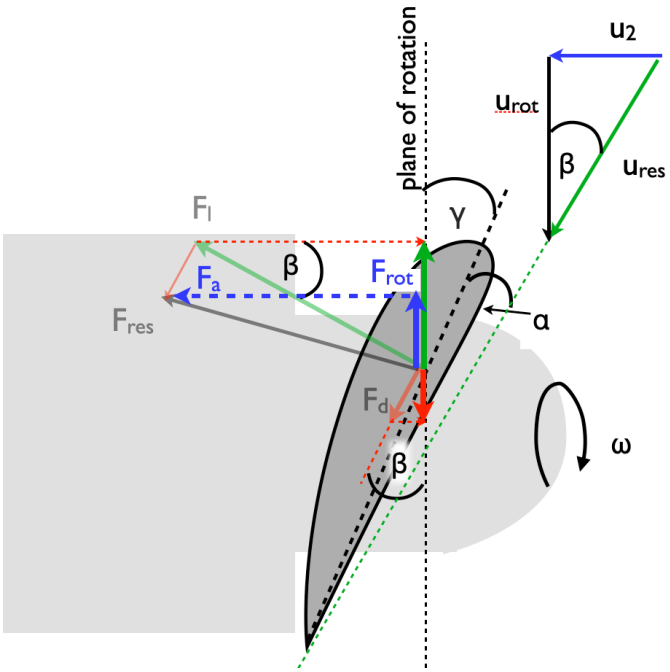

Fig. 2.7 : (color online) Cut through an airfoil rotating around the wind turbine axis. The rotational velocity $u_{rot} = \omega r$ (black) is perpendicular to the axial velocity $u_2 = 2u_1/3$ (blue). The angle of attack α is the angle between the air velocity u_{res} (green) and airfoil chord (dashed black line). The angle β denotes the angle between the air velocity u_{res} and the plane of rotation. γ denotes the angle between the chord and the plane of rotation. The lift and drag forces F_l and F_d are displayed (green and red), giving the total force F (gray), as well as the rotational and thrust projections F_{rot} and F_a (blue).

thrust. Considering an infinitesimal cut dr at radial position r in the polar plane of the rotor, the infinitesimal components are

$$\begin{aligned} dF_{rot} &= \frac{\rho}{2} u_{res}^2 \; c(r) \; dr \Big[C_l(\alpha) \; \sin(\beta) - C_d(\alpha) \; \cos(\beta) \Big] \,, \\ dF_a &= \frac{\rho}{2} u_{res}^2 \; c(r) \; dr \Big[C_l(\alpha) \; \cos(\beta) + C_d(\alpha) \; \sin(\beta) \Big] \,. \end{aligned} \tag{2.28}$$

Now we take the chord as a quantity that may change with the radius, i.e. $c = c(r)$.

Involved angles

The angle β from the velocity triangle (see figure 2.7) reads

$$\tan(\beta) \;=\; \frac{2u_1/3}{\omega r} \;=\; \frac{2}{3}\frac{u_1}{\omega R}\frac{R}{r} \;=\; \frac{2}{3}\frac{R}{r\lambda} \,, \tag{2.29}$$

and denotes the angle between the resulting velocity u_{res} and the plane of rotation at the distance to the root r. γ describes the local angle between the profile's chord and the plane of rotation and is a result of the designed twist of the rotor blade, where α is the local angle of attack of the profile with respect to the resulting velocity u_{res}. The angles are connected through

$$\alpha \;=\; \beta - \gamma \,. \tag{2.30}$$

Pitching the whole rotor blade by the angle θ will therefore change the orientation γ of the profile allowing to actively control the aerodynamic forces through the angle of attack α.

Power of blade section dr

Only the rotational component F_{rot} is of use to rotate the rotor. The force F_a in the axial direction does not contribute to the power production but to the thrust acting on the turbine structure. F_a should be minimized to reduce mechanical fatigue. The infinitesimal power (force times velocity) associated to the rotational force acting on z rotor blades is

$$dP_{rot} \;=\; z\; dF_{rot}\; \omega r \;=\; z\; \frac{\rho}{2} u_{res}^2\; c(r)\;\; dr \Big[C_l \sin(\beta) - C_d \cos(\beta) \Big]\; \omega r\,. \tag{2.31}$$

Optimized power leads to $c(r)$

The goal is to construct the blades in such a way that they extract the optimal power following Betz limit out of the wind. It should be noted that in this idealized case, the drag force is taken to be zero, which can be practically approached if $C_d \ll C_l$. Each infinitesimal radial annulus of size $2\pi r dr$ should extract a fraction $16/27$ of the wind power

$$dP_{ideal} \;=\; dP_{rot} \quad\Longleftrightarrow\quad \frac{16}{27}\,\frac{\rho}{2}\, u_1^3\; (2\pi r dr) \;=\; z \frac{\rho}{2} u_{res}^2\; c(r)\; \omega r\; dr\; C_l \sin(\beta) \tag{2.32}$$

based on equation (2.31). In order to approach such ideal case, the blade depth $c(r)$ must follow

$$c(r) \;=\; \frac{16\pi}{9}\,\frac{1}{z\; C_l}\,\frac{R}{\lambda}\,\frac{1}{\sqrt{\lambda^2 \frac{r^2}{R^2} + \frac{4}{9}}}\,. \tag{2.33}$$

For $r > R/7$ the square root gets approximately $\lambda \frac{r}{R}$ and thus we obtain

$$\boxed{\frac{c(r)}{R} \;\approx\; \frac{16\pi}{9}\,\frac{1}{z C_l \frac{r}{R} \lambda^2}}\,. \tag{2.34}$$

Here all lengths are given with respect to the blade length R. This approximation has an important consequence on the design of rotor blades. The depth $c(r)$ decreases when increasing either the number of blades, z, the lift coefficient C_l, the radius r or the tip speed ratio λ. This explains why fast rotating wind turbines tend to have only two or three narrow blades to optimize power extraction, while old western-mill machines have many, rather broad blades (in order to maximize mechanical torque).

As a last remark we point out that the optimized chord function $c(r)$ of eq. (2.34) was derived without any condition on the value of $C_l(\alpha)$. Thus an optimal blade shape can be designed also for values of the angle of attack α below the one that gives the maximal lift force. It is good to avoid maximal lift forces as for such conditions small fluctuations in the angle of attack may lead to the stall effect, a sudden decrease in lift.

2.3 IEC power curve and annual energy production

In the previous sections, we have set theoretical models to predict and optimize power production, however, with 'laminar flow' approaches. Now, we look at real wind turbines, that operate in a ***turbulent, non stationary wind***. We describe a normalized way of ***monitoring power production performances***. We assume that wind velocities u and power P are measured at a sampling frequency of the order of a few Hz, and focus on the ***processing*** of these data. This processing is mainly performed in two steps.

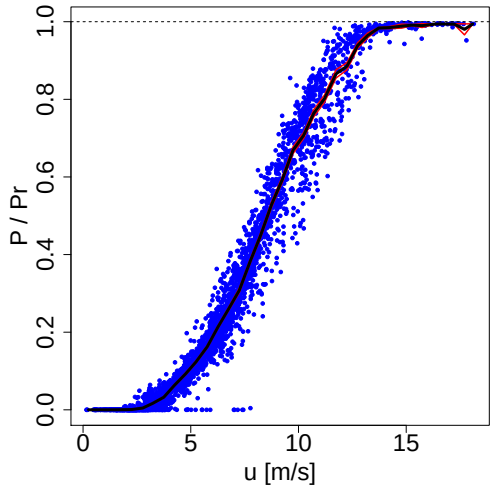

Fig. 2.8 : (color online) Typical power curve (black line) and corresponding error bars (thin red lines underneath) obtained according to the IEC norm. The blue dots represent the 10-minute average values.

After adequate normalization of the data, the first step consists in averaging the measured data over time intervals of 10 minutes. The ***IEC power curve*** (IEC 2005) is derived in a second step from the ten-minute averages using the so-called method of bins, i.e. the data is separated into wind speed intervals of width 0.5 m/s. In each of these intervals, labeled i, bin averages of wind speed u_i and power output P_i are calculated according to

$$u_i = \frac{1}{N_i}\sum_{j=1}^{N_i} u_{norm,i,j} \quad , \quad P_i = \frac{1}{N_i}\sum_{j=1}^{N_i} P_{norm,i,j}\,, \tag{2.35}$$

where $u_{norm,i,j}$ and $P_{norm,i,j}$ are the normalized 10-minute average values of wind speed and power, and N_i is the number of 10 min data sets in the i^{th} bin.

For the power curve to be complete and reliable, each wind speed bin must include at least 30 minutes of sampled data. Also, the total measurement time must cover at least a period of 180 hours. The range of wind speeds must range from 1 m/s below cut-in wind speed to 1.5 times the wind speed at 85% of the rated power of the wind turbine. The norm also provides an estimation of uncertainty as the standard error of the normalized power data, plus additional uncertainties related to the instruments, the data acquisition system and the surrounding terrain. A typical IEC power curve is presented in figure 2.8.

The IEC norm also defines the ***annual energy production*** or AEP, as it will be presented in section 2.3.2 below. The AEP is a central feature for economical considerations, as it gives a first estimate of the long-time energy production of a wind turbine. As it sets a unique ground for wind power performance worldwide, the IEC norm helps building a general understanding between manufacturers, scientists and end-users. This statement comes to be ever more important as the wind energy sector grows. Hence, focusing on this standard is paramount to any study on power performance.

2.3.1 Turbulence-induced deviations

As a downside to its simplicity, the IEC power curve method presents a limitation. It suffers a physical and mathematical imperfection. In order to deal with the complexity of the wind speed

and power signals, the data is systematically averaged over time. Although a statistical averaging is necessary to extract the main features from the complex processes, the averaging procedure over 10-minute intervals lacks a clear physical meaning, beyond its statistical definition. As the wind[3] fluctuates on various time scales (down to seconds and less), a systematic averaging over ten minutes filters out all the short-scale turbulent dynamics. Combining these turbulent fluctuations with the nonlinear power curve $P(u)$, $P(u)$ being proportional to u^3 for small u, and being more complicated for large u, the resulting IEC power curve is spoiled by mathematical errors. To show this, one can first split the wind speed $u(t)$ sampled at 1Hz into its mean value and the fluctuations around this mean value

$$u(t) = \overline{u(t)} + u'(t)\,, \tag{2.36}$$

where the operation $\overline{x(t)}$ on a given signal $x(t)$ represents the 10-minutes average of $x(t)$ as defined by the IEC norm. Assuming that $u'(t) \ll \overline{u(t)}$, a Taylor expansion of $P(u(t))$ reads (Böttcher *et al.* 2007)

$$\begin{aligned} P\left(u(t)\right) &= P\left(\overline{u(t)}\right) + u'(t)\left(\frac{\partial P(u)}{\partial u}\right)_{u=\overline{u(t)}} \\ &\quad + \frac{u'(t)^2}{2!}\left(\frac{\partial^2 P(u)}{\partial u^2}\right)_{u=\overline{u(t)}} + \frac{u'(t)^3}{3!}\left(\frac{\partial^3 P(u)}{\partial u^3}\right)_{u=\overline{u(t)}} + o\left(u'(t)^4\right)\,. \end{aligned} \tag{2.37}$$

Averaging equation (2.37) yields

$$\begin{aligned} \overline{P\left(u(t)\right)} &= P\left(\overline{u(t)}\right) + 0 \\ &\quad + \frac{\overline{u'(t)^2}}{2}\left(\frac{\partial^2 P(u)}{\partial u^2}\right)_{u=\overline{u(t)}} + \frac{\overline{u'(t)^3}}{6}\left(\frac{\partial^3 P(u)}{\partial u^3}\right)_{u=\overline{u(t)}} + o\left(u'(t)^4\right)\,, \end{aligned} \tag{2.38}$$

because $\overline{u'(t)} = \overline{u(t) - \overline{u(t)}} = 0$. This means that the average of the power is not equal to the power of the average, and must be corrected by the 2^{nd} and 3^{rd}-order terms. As the IEC power curve directly relates the 10-minute averages of wind speed and of power output, it neglects the higher-order terms in the Taylor expansion. The 2^{nd}-order term is the product of the variance $\sigma^2 = \overline{u'(t)^2}$ of u(t)[4] and the second-order derivative of the power curve[5]. This demonstrates that the IEC power curve cannot describe in a mathematically rigorous way the nonlinear relation of power to wind speed when coupled with wind fluctuations (stemming from turbulence), at least not without higher-order corrections.

As a consequence of this mathematical over-simplification, the result depends on the 'turbulence intensity' $I = \sigma/\overline{u}$, so on the wind condition during the measurement Böttcher *et al.* (2007). It is illustrated in figure 2.9, where the IEC power curve deviates from the theoretical power curve with increasing turbulence intensity, as predicted by equation (2.38). As it does not characterize the wind turbine only, but also the measurement condition, this raises the question of its reproducibility and stability.

[3] To some extent the power output also fluctuates on short time scales, but its high-frequency dynamics are limited by the inertia of the wind turbine.

[4] $\sigma^2 = \overline{(u - \overline{u})^2} = \overline{u'(t)^2}$.

[5] Assuming a cubic power curve $P(u) \propto u^3$, $P(u)$ has non-zero derivatives up to 3^{rd}-order. Moreover, the transition point to rated power may have non-zero derivatives of arbitrary order, see figure 2.9.

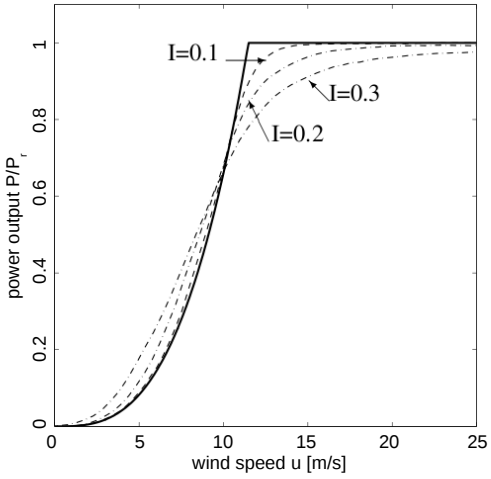

Fig. 2.9 : Typical IEC power curves for various turbulence intensities $I = 0.1,\ 0.2,\ 0.3$ (dashed lines). The full line represents the theoretical power curve. This result was obtained from numerical model simulations from Böttcher *et al.* (2007).

2.3.2 Annual energy production

The one-dimensional *limitation* of the IEC power curve becomes an advantage for long-term energy production, as $P_{IEC}(u)$ relates unambiguously a unique value of power for each wind speed. As the AEP estimates the energy produced over a year, it can be seen as a prediction estimate. A prediction of power production at high-frequency is also possible using the Langevin approach, as it will be shown in the next section. The estimation of the AEP extrapolates the power production of a wind turbine characterized by its power curve in a given location. Here we do not give an exact transcription of the AEP procedure from the IEC norm (IEC 2005), but rather a comprehensive introduction on how power production can be estimated simply from a wind speed measurement. For such, the *AEP procedure* introduced here is not the official AEP procedure following IEC, but a similar version. In both cases, the availability of the wind turbine is assumed to be 100%.

Estimating the wind resource

Any location scheduled to host a wind turbine can be categorized in advance by a characterization of its wind resource. A local measurement of wind speed from a met mast at hub height[6] of the hypothetical wind turbine must be performed, typically over one year[7]. From this wind speed measurement $u(t)$, a ten-minute (or hourly) averaging is applied on $u(t)$. The probability density function (PDF) $f(u_i)$ of the ten-minute average values u_i is established. For clarity, the values u_i will be labelled u. $f(u)$ returns the probability of occurrence of the wind speed u. For long enough measurements, $f(u)$ is known to fit a ***Weibull distribution*** (Richardson 1922)

$$f(u;\lambda,k) = \frac{k}{\lambda}\left(\frac{u}{\lambda}\right)^{k-1} e^{(-u/\lambda)^k}\,, \tag{2.39}$$

[6]Typical hub heights of commercial multi-MW class wind turbines are in the order of 100 m, justifying the interest for a portable LIDAR sensor...

[7]A measurement of wind speed over one year covers the various wind situations resulting from various seasonal behaviors.

where k and λ[8] are called respectively the ***shape*** and ***scale factors*** [9]. Visual examples of such wind speed distributions are given in Burton *et al.* (2001).

Estimating the AEP

A given wind site is characterized for the AEP by its wind speed PDF $f(u)$, while a given wind turbine is characterized by its IEC power curve. As P_{IEC} relates unambiguously a given wind speed u to the corresponding average power output $P_{IEC}(u)$, the power curve serves as a transfer function from wind speed to average power output. An estimation of the average power output $\overline{P}$ can be obtained following

$$\overline{P} \;=\; \int_0^\infty f(u)\; P_{IEC}(u)\, du\,, \tag{2.40}$$

and an estimation for the energy production over a period T reads

$$T\;\overline{P} \;=\; T\int_0^\infty f(u)\; P_{IEC}(u)\, du\,. \tag{2.41}$$

Over one year, $T = 8766$ hours, therefore

$$AEP \;=\; \overline{P}\;8766\,, \tag{2.42}$$

where $\overline{P}$ is given in Watt and AEP is given in Watt hour.

Thanks to its simple mathematical procedure, the AEP is commonly used to make rough predictions of energy production, as well as for financial estimations. It can predict how much energy a wind turbine will generate on a given site before installing it. This allows for an optimal choice of design for the optimal location. This result however remains a rough estimation, as it neglects e.g. wake losses generated by other surrounding wind turbines.

2.4 A new alternative: the Langevin power curve

An alternative to the standard IEC power curve is proposed in this section. As the IEC norm defines the measurement procedure with relevance, the same conditions will be considered for the Langevin analysis. The difference lies in the different approach to process the measured data.

One additional point on the sampling frequency is however important for the Langevin analysis. Because the method resolves the dynamics of a wind turbine in the order of seconds, a minimum sampling frequency in the order of 1Hz is necessary for the measurements of wind speed and power output.

[8] λ is not the tip speed ratio of a wind turbine, but a parameter of the Weibull distribution.

[9] IEC (2005) refers to the Rayleigh distribution, which is a special case of the Weibull distribution for $k = 2$.

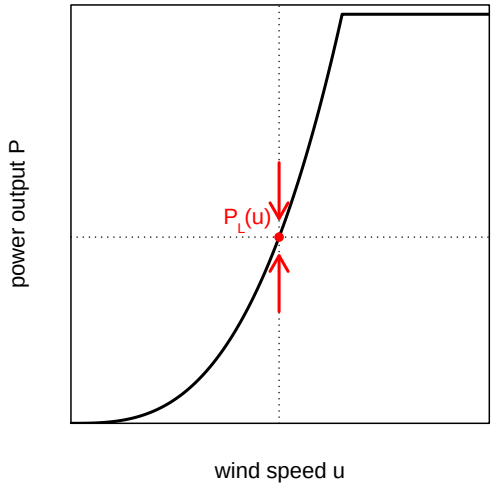

Fig. 2.10 : Illustration of the concept of stable fixed point $P_L(u)$. For constant wind speed, the power output would relax to a stable value $P_L(u)$. This sketch is inspired from Anahua *et al.* (2008).

2.4.1 A dynamical concept

The power characteristic of a wind turbine can be derived from high-frequency measurements without using temporal averaging. One can regard the power conversion as a relaxation process which is driven by the turbulent wind fluctuations (Rosen & Sheinman 1994; Rauh & Peinke 2004). More precisely, the wind turbine is seen as a dynamical system which permanently tries to adapt its power output to the fluctuating wind. For the (hypothetical) case of a laminar inflow at constant speed u, the power output would relax to a fixed value $P_L(u)$[10], as illustrated in figure 2.10. Mathematically, these *attractive* power values $P_L(u)$ are called *stable fixed points* of the power conversion process.

2.4.2 The Langevin equation

The Langevin power curve[11] is derived from high-frequency measurements of wind speed $u(t)$ and power output $P(t)$. All necessary corrections and normalizations from the IEC norm IEC (2005) should be applied on the two time series.

The wind speed measurements are divided into bins u_i of 0.5 m/s width, as done in IEC (2005). This accounts, to some degree, for the non-stationary nature of the wind, yielding quasi-stationary segments $P_i(t)$ for those times t with $u(t) \in u_i$. The following mathematical analysis will be performed on these segments $P_i(t)$. From now on, the subscript i will be omitted and the term $P(t)$ will refer to the quasi-stationary segments $P_i(t)$. The power conversion process is then modeled by a first-order stochastic differential equation called the ***Langevin equation***[12]

$$\boxed{\frac{d}{dt}P(t) \;=\; D^{(1)}(P) \;+\; \sqrt{D^{(2)}(P)}\,\Gamma(t)}\,. \tag{2.43}$$

[10]The subscript L stands for "Langevin" as $P_L(u)$ will be associated to the formalism of the Langevin equation.

[11]In former publications on the topic, the Langevin power curve was called *dynamical power curve* or *Markovian power curve*. It is nonetheless the same approach.

[12]This equation is the reason for the name of the Langevin power curve.

In this model, the time evolution of the power output is controlled by two terms[13].

- $D^{(1)}(P)$ represents the deterministic relaxation of the wind turbine, leading the power output towards the attractive fixed point $P_L(u)$ of the system. For such, $D^{(1)}(P)$ is commonly called the ***deterministic drift function***.
- The second term $\sqrt{D^{(2)}(P)}\,\Gamma(t)$ represents the stochastic (random) part of the time evolution, and serves as a simplified model for the turbulent wind fluctuations that drive the system out of equilibrium. The function $\Gamma(t)$ is a Gaussian-distributed, delta-correlated ***noise*** with variance 2 and mean value 0. $D^{(2)}(P)$ is commonly called the ***diffusion function***. A mathematical approach to the Langevin equation can be found in Risken (1996).

The mathematical background for this 'Langevin' approach is presented in the Appendix A, which is an introduction to stochastic theory.

2.4.3 The drift function and the Langevin power curve

The deterministic drift function $D^{(1)}(P)$ is of interest as it quantifies the relaxation of the power output towards the stable fixed points of the system. When the system is in a stable state, no deterministic drift occurs[14], and $D^{(1)}(P) = 0$. Following equation (2.44), $D^{(1)}(P)$ can be understood as the average time derivative of the power signal $P(t)$ in each region of wind speed u_i and power output P.

The drift and diffusion functions can be derived directly from measurement data as conditional moments (Risken 1996)

$$D^{(n)}(P) = \lim_{\tau \to 0} \frac{1}{n!\tau} \langle\, (P(t+\tau) - P(t))^n \mid P(t) = P \,\rangle_t \,, \tag{2.44}$$

where $n = 1, 2$ respectively for the drift and diffusion functions. The averaging $\langle \cdot \rangle_t$ is performed over t, as the condition means that the calculation is only considered for those times during which $P(t) = P$.

This means that the averaging is done separately for each wind speed bin u_i and also for each level of the power P. One could speak of a *state-based* averaging on u and P, in contrast to the temporal averaging performed in the IEC norm. A typical drift function is displayed in figure 2.11.

The dynamics of the power signal can be directly related to the local sign and value of $D^{(1)}$. A positive drift indicates that the power tends to increase (arrows pointing up in figure 2.11), in regions where the wind turbine does not produce enough power for the given wind speed. On the contrary, a negative drift corresponds to a decreasing power (arrows pointing down), in regions where the wind turbine produces too much power for the given wind speed. At the intersection are the points where $D^{(1)} = 0$, indicating that when at this value, the power

[13] $D^{(1)}$ and $D^{(2)}$ are the first two coefficients of the general Kramers-Moyal coefficients.

[14] To separate stable (attractive) from unstable (repulsive) fixed points, also the slope of $D^{(1)}(P)$ must be considered.

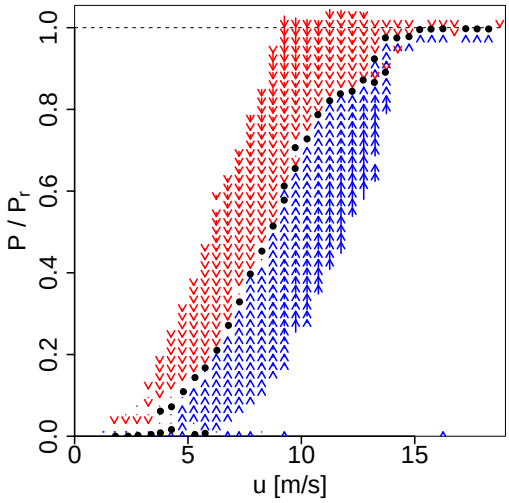

Fig. 2.11 : Typical drift function $D^{(1)}(P)$. Each arrow represents the local value of $D^{(1)}(P)$ in magnitude (length of the arrow) and direction (pointing up for positive values). The stable fixed points where $D^{(1)}(P) = 0$ are given by the black dots.

output is in a stable configuration (the average time derivative is zero). The collection of all the points where the drift function is zero is defined as the ***Langevin power curve***, and will be further labelled $P_L(u)$.

The stable fixed points $P_L(u)$ of the power conversion process can be extracted from the measurement data as solutions of

$$D^{(1)}\left(P_L(u)\right) = 0\,. \tag{2.45}$$

An illustration is given in figure 2.12.

Following the mathematical framework of equations (2.43) and (2.44), an estimation of uncertainty for $P_L(u)$ can be performed (Gottschall 2009). One can see on figure 2.12 that, for most wind speeds, the power curve has very little uncertainty. Nevertheless, larger uncertainties occur in the region of transition to rated power. There the power conversion is close to stability over a wider range of power values, as a consequence of the changing control strategy from partial load to full load operation (see figure 2.6). It is a region of great interest as the controller of the wind turbine is highly solicited for the transition to rated power.

2.4.4 Advantages of the Langevin approach

The Langevin equation (2.43) is a simplifying model for the power conversion process. The question of its validity for wind turbine power signals was positively answered in recent developments (Milan *et al.* 2010), as the power signal of a wind turbine could be successfully modelled. Also, the drift function $D^{(1)}$ is well-defined for a large class of stochastic processes, and is not limited only to the class of the Langevin processes.

Moreover, the definition of the drift function does not suffer the systematic errors caused by temporal averaging. For such, the Langevin power curve characterizes the wind turbine dynamics only, regardless of the wind condition during the measurement[15]. The results are therefore machine-

[15] Assuming that the measurement period is sufficiently long to reach statistical convergence.

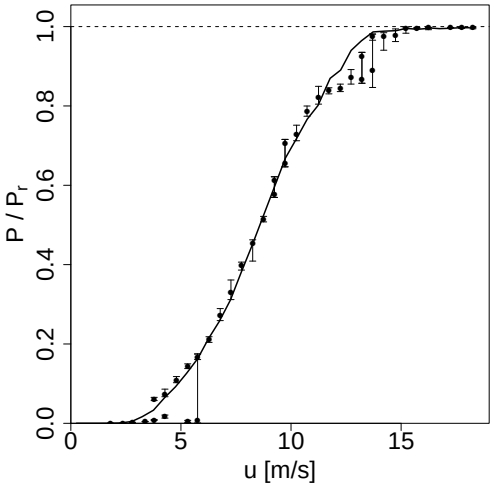

Fig. 2.12 : Typical Langevin power curve (black dots with corresponding error bars) and IEC power curve (solid line).

dependent only, and not site- or measurement-dependent, as the intensity of turbulence has no influence on the Langevin power curve.

Additionally, this approach can show complex characteristics of the investigated system, such as regions where the system is close to stability, as mentioned above, or multiple stable states, see also Anahua *et al.* (2008); Gottschall & Peinke (2008). For these various reasons, the Langevin power curve represents a promising tool for power performance monitoring, too.

2.4.5 Detecting dynamical anomalies

The intention of ***dynamical monitoring*** is to detect dynamical anomalies that appear on operating wind turbines. Monitoring refers to the time evolution of the power performance here. A good monitoring procedure should be reliable[16], as fast as available, and possibly also inform on the source of the anomaly. While monitoring procedures come to be ever more complex, the approach presented here is based only on a power curve estimation. This approach is not intended to give a full-featured method, but rather an illustration of the amount of information given by power curves. More advanced studies on the topic of power curves for monitoring are being developed, but remain outside the scope of this introduction as they represent active research topics.

The monitoring procedure simply consists in computing $P_L(u)$ at an initial time that will serve as a reference[17]. Potential changes in time of $P_L(u)$ are considered anomalies, or malfunctions inside the wind turbine that spoil the conversion dynamics. While this strategy is very simple, the challenge lies in defining the right threshold for a change in $P_L(u)$ to be considered an anomaly. This threshold, along with other parameters such as the necessary measurement time or time reactivity of the method depend on the wind turbine design and location.

To illustrate the ability of the method, the monitoring procedure was applied on a numerical simulation. The simulation was applied on measurement data, where an anomaly was introduced. This artificial anomaly limits the power production to $P \simeq 0.55\, P_r$ for intermediate wind speeds,

[16] An over-sensitive procedure might indicate non-existing anomalies, while an under-sensitive procedure would fail to detect a major malfunction.

[17] The reference time is chosen when the wind turbine is believed to work with full capacity.

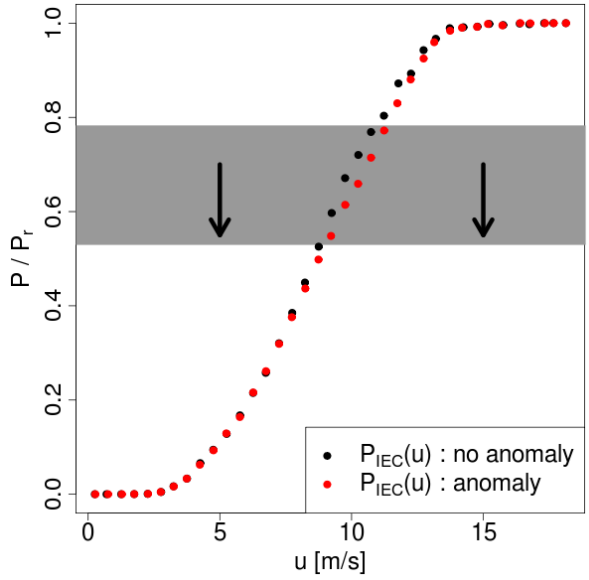

Fig. 2.13 : (color online) Comparison of $P_{IEC}(u)$ before the anomaly (black in background) and after the anomaly (grey or red in front). The artificial anomaly was applied in the grey rectangle.

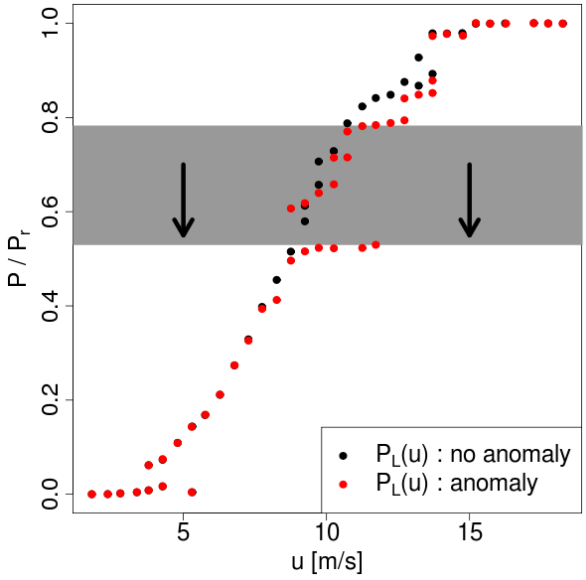

Fig. 2.14 : (color online) Comparison of $P_L(u)$ before the anomaly (black in background) and after the anomaly (grey or red in front). The artificial anomaly was applied in the grey rectangle.

as represented by the grey rectangles in figure 2.13 and 2.14. More clearly, when in this rectangle, the power signal was sometimes forced to reduce towards 0.55 P_r. From this artificial data, $P_L(u)$ and $P_{IEC}(u)$ were then computed and compared to the original data. This is illustrated in figure 2.13 and 2.14.

Similar anomalies were observed on several real wind turbines, justifying the reason for this artificial anomaly. For information, the total energy production was reduced to 96.6% compared to the original energy production due to the presence of the anomaly. figure 2.14 illustrates the higher reactivity of $P_L(u)$. While in figure 2.13 $P_{IEC}(u)$ only shows a minor deviation in the region of the anomaly, $P_L(u)$ clearly *detaches* from the typical cubic curve to adjust to the new dynamics. $P_L(u)$ can detect changes in the dynamics of the conversion process, unlike $P_{IEC}(u)$ that is better suited for the AEP.

The Langevin power curve is more reactive to changes in the dynamics. As the IEC power curve averages over 10 minutes intervals, the information about high-frequency dynamics is lost. Also, the second averaging in wind speed prevents from seeing multi-stable behaviors.

In addition, the Langevin power curve does not depend on the turbulence intensity, unlike the IEC power curve, see section 2.3.1. A deviation in the Langevin power curve indicates a change in the conversion dynamics, regardless of the wind situation. This makes the Langevin power curve a promising tool for dynamical monitoring.

2.5 Turbulence modelling

For the bibliography, especially concerning the 'old' papers, see Frisch (1995).

2.5.1 General definitions - Experimental results

The question here, to go beyond the basic description in terms of scales introduced in Plaut (2021*a*), is how to characterize the ***spatial complexity of a turbulent field***. Of course, the starting point is still the definition of the two length scales between which the complex structures of turbulence are. There is a large scale L, called ***integral length scale***[18]. For scales larger than L, if they exist[19], unstructured, i.e. uncorrelated random fluctuations are left. In other words, L can often be defined as a correlation length. There is a small scale η, called ***dissipation or Kolmogorov scale***[20]. Any turbulent structure is smooth out by dissipation for length smaller then η. Thus the complex structure of turbulence lives at length scales between these two limiting scales L and η.

Besides these length there is the ***'turbulent energy'*** which determines the complexity. As turbulence is a dissipative structure, it can be sustained only if there is a permanent flow of energy into the system, i.e., a permanent power driving the flow structure. For a steady state there must be an equilibrium between the driving power and the dissipated power. This size is measured as a quantity ϵ called turbulent energy, what is not precise, as it is a power. Normalized by the mass the dimension is $[\mathrm{m}^2/\mathrm{s}^3]$. The cascade picture of turbulence is that this power ϵ is fed into the system at large scales, transfered by a cascade process to smaller scales and dissipated at smallest scales.

The spatial dependencies can be investigated by studying the ***velocity increments*** and their ***structure functions***, as defined below for a length scale r,

$$\eta \leq r \leq L \ .$$

Hereafter u (resp. $\mathbf{u}$) denotes one component of (resp. the whole) velocity field. One uses a ***Reynolds decomposition***

$$\boxed{u \ = \ \langle u \rangle \ + \ u'} \tag{2.46}$$

with u' the fluctuating velocity.

Definitions

Velocity increments:

$$\boxed{u_r(x) := u(x+r) - u(x)} \ . \tag{2.47}$$

Structure functions:

$$\boxed{S^n(r) := \langle (u_r(x))^n \rangle = \langle (u(x+r) - u(x))^n \rangle} \ . \tag{2.48}$$

[18] Denoted ℓ in Plaut (2021*a*).

[19] Often, L can be the size of the system.

[20] Denoted ℓ_K in Plaut (2021*a*).

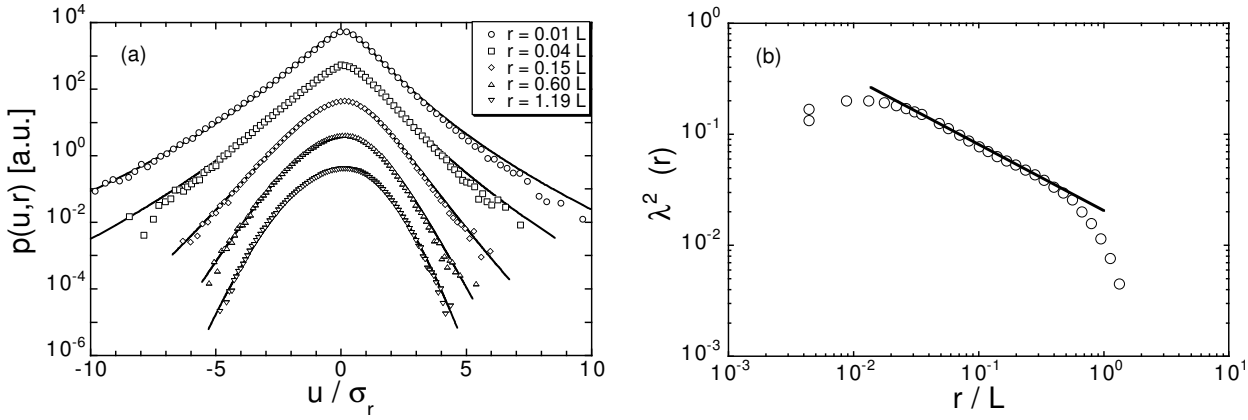

Fig. 2.15 : Statistical analysis of Experimental data acquired with Hot Wire Anemometry on a Laboratory Turbulent Flow, an air into air round free jet (Renner *et al.* 2001; these data are also available on the web page of the module). (a) : PDF $p(u_r)$ of the velocity increments u_r for different values of r; the values of u_r are normalized with the corresponding standard deviation σ_r of u_r. (b) : Corresponding form parameters, see equation (2.71) below.

Comments

In principle all can be expressed for vectors: $u \to \mathbf{u}$, $x \to \mathbf{x}$ and $r \to \mathbf{r}$.

Some ***probability density functions*** (PDF) $p(u_r)$ of velocity increments u_r for different values of r are shown in figure 2.15a. One observes a non-Gaussian character for small values of r, with events corresponding to large increments that are rather 'frequent', at least, with respect to what a Gaussian PDF would give. This is a signature of the ***intermittent character*** of small-scale turbulent flows. Observe that, for larger values of r, the PDF become more and more Gaussian. We will write a model for $p(u_r)$ in section 2.5.4, with the 'multiplicative cascade approach' by Castaing. However, firstly, we present 'older' models that paved the way to arrive to this point of view.

2.5.2 Kolmogorov & Obukov 1941

From the idea of a ***cascade***, Kolmogorov deduced that the structure function should be $S^n(r) = f(\epsilon, r)$ with ϵ the energy (power density) transferred in the cascade. Using dimensional analysis, one gets

$$S^n(r) \;=\; f(\epsilon, r) \;=\; C_n\, \epsilon^{n/3}\, r^{n/3} \tag{2.49}$$

with C_n a universal constant. For $n = 2$ we obtain $S^2(r) \sim r^{2/3}$, and for $n = 3$, $S^3(r) \sim r$, as tested experimentally in figure 2.16. Note the dimension of the "transferred energy" in the cascade $\epsilon \;\equiv\; E/(t\; m) \;\equiv\; \ell^2/t^3$.

Karmann, Howarth & Kolmogorov derived from the Navier-Stokes equation and isotropy the so-called $-\frac{4}{5}$***-law***

$$S^3(r) = -\frac{4}{5}\epsilon r + 6\nu \frac{dS^2(r)}{dr}\ . \tag{2.50}$$

According to the equation (2.49) $6\nu\frac{dS^2(r)}{dr} \propto r^{-1/3}$, thus for large scales ($r \to L$) one has $S^3(L) \propto$

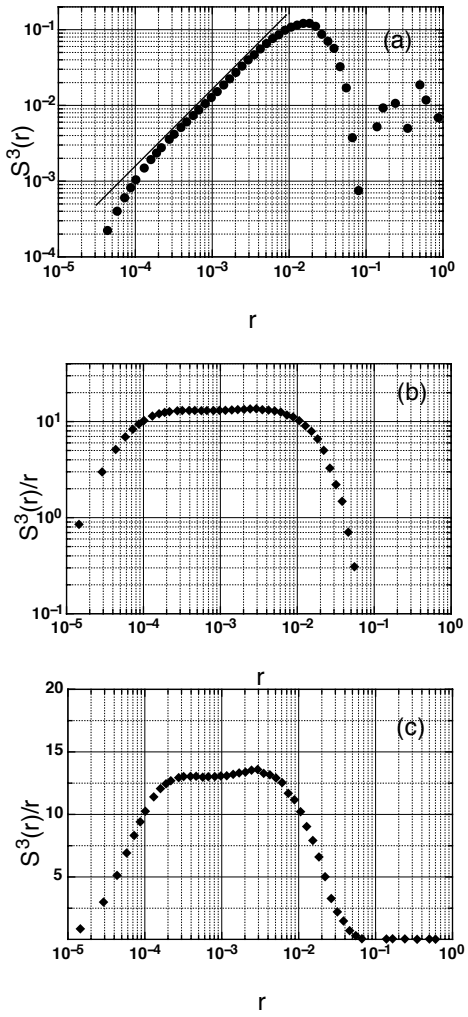

Fig. 2.16 : Experimental test of the law $|S^3| \propto r$, by Chabaud, B. & Chanal, O., CNRS Grenoble, France; reprinted from Friedrich & Peinke (2009). **(a)** Absolute value of the third-order structure function plotted with log-log scales. **(b)** and **(c)** absolute value of the compensated third-order structure function plotted with **(b)** log-log scales **(c)** log-lin scales.

$-\frac{4}{5}\epsilon L$. From this L/η can be estimated. Using

$$\epsilon \approx \frac{S^3}{L} = \frac{< (u(x+L) - u(x))^3 >}{L} \approx \frac{< u'^3 >}{L} , \tag{2.51}$$

one obtains

$$\boxed{L/\eta = \frac{L}{\left(\frac{\nu^3}{\epsilon}\right)^{1/4}} = \frac{L\epsilon^{1/4}}{\nu^{3/4}} \approx \frac{L\frac{<u'^3>}{L}^{1/4}}{\nu^{3/4}} = \left(\frac{u'L}{\nu}\right)^{3/4} = Re^{3/4}} . \tag{2.52}$$

2.5.3 Kolmogorov & Obukov 1962

L.D. Landau pointed out that

> *'It is not obvious why ϵ is not a fluctuating quantity'.*

The idea of Kolmogorov 1941 can be taken as $\langle \epsilon_L \rangle = \langle \epsilon_r \rangle = \langle \epsilon_\eta \rangle$ - i.e the mean transfered energy in the cascade is conserved.

After the comment of Landau, Kolmogorov claimed that it is reasonable for $\epsilon(r)$ to assume a ***log-normal distribution***

$$p(\epsilon) \;=\; \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\ln x)^2}{2\sigma^2} \right) \;. \tag{2.53}$$

Short argumentation for log-normal distribution

Some notations: The length scales of the turbulent cascade are denoted as $r_n < r_{n-1} < ...r_0 = L$, and $\epsilon_{r_i} := \epsilon_i$ the energy transferred at scale r_i.

The idea of a cascade is that now the sequence of $\epsilon_{r_1} = \epsilon_{r_2} = \ldots = \epsilon_L$ may become random by multipliers such that

$$\epsilon_i \;=\; a_i \; \epsilon_{i-1} \;.$$

The energy conservation is given by the condition $< a_i > \;=\; 1$. Thus one gets for the cascade:

$$\epsilon_{r_n} \;=\; a_n \; \epsilon_{r_{n-1}} \;=\; a_n \; a_{n-1} \epsilon_{r_{n-2}} \;=\; \ldots \;=\; a_n \; a_{n-1} \; \ldots \; a_1 \; \epsilon_L \;. \tag{2.54}$$

Taking the log of this equation:

$$\ln \frac{\epsilon_{r_n}}{\epsilon_L} = \sum_{i=1}^{n} \ln a_i \;. \tag{2.55}$$

This is now taken as a sum over independent random numbers $\ln a_i$. The "central limit theorem" from Kolmogorov says that such a sum of independent random numbers converges towards a Gaussian distribution, hence

$$p(\epsilon_{r_n}) \;=\; p\left(\ln \frac{\epsilon_{r_n}}{\epsilon_L} \right) \;=\; \frac{1}{\sqrt{2\pi\Lambda^2}} \exp\left(-\frac{\left(\ln \frac{\epsilon_{r_n}}{\epsilon_L} \right)^2}{2\Lambda^2} \right) \tag{2.56}$$

where the variance Λ^2 can still be a function of r_n.

Kolmogorov 1962 assumes that

$$\Lambda^2(r) \;=\; \Lambda_0^2 - \mu \ln \frac{r}{L} \;. \tag{2.57}$$

Some argumentation for Kolmogorov 1962 hypothesis

As $\ln a_i$ are assumed to be uncorrelated ($\langle (\ln a_i)(\ln a_j) \rangle = 0$),

$$\Lambda^2(r) \;=\; \left\langle \left(\ln \frac{\epsilon_r}{\epsilon_L} \right)^2 \right\rangle \;=\; \left\langle \left(\sum_{i=1}^{n} \ln a_i \right)^2 \right\rangle \;=\; n \left\langle (\ln a_i)^2 \right\rangle \;\propto\; n \;. \tag{2.58}$$

Thus $\Lambda^2(r) \sim n$ with n the depth of the cascade. Furthermore it is assumed that $r_{n+1} = k\, r_n$, for example $k = \frac{1}{2} \quad \rightarrow \quad r_{n+1} = \frac{1}{2}\, r_n$. Thus

$$\begin{aligned} r_{n+1} = k\, r_n \quad &\Longleftrightarrow \quad r_n = k^n\, L \\ &\Longleftrightarrow \quad k^n = \frac{r_n}{L} \\ &\Longleftrightarrow \quad n\, \ln k = \ln \frac{r_n}{L} \\ &\Longleftrightarrow \quad n = \ln \frac{r_n}{L}\, \frac{1}{\ln k}\ . \end{aligned} \tag{2.59}$$

As $k < 1$, is follows that $\ln k < 0$. Thus we can define $\mu := -\frac{1}{\ln k}$ and one obtains

$$n = -\mu \ln \frac{r}{L}\ . \tag{2.60}$$

Thus the hypothesis follows

$$\Lambda^2(r) = \Lambda_0^2 - \mu \ln \frac{r}{L} \tag{2.61}$$

with μ an ***intermittency coefficient***.

Knowing $p(\epsilon)$ and $\Lambda^2(r)$ one can calculate

$$\langle \epsilon^{n/3} \rangle \sim r^{-\mu \frac{n(n-3)}{18}}\ .$$

This is the results of Obukov & Kolmogorov 1962.
One deduces from this that the structure functions

$$S^n(r) = C_n \langle \epsilon^{n/3} \rangle r^{n/3} \sim C_{\epsilon_n} r^{n/3 - \mu \frac{n(n-3)}{18}} \tag{2.62}$$

This equation is known as the ***intermittency correction*** to Kolomogov 1941 (2.49). It is easy to see that for $n = 3$ (2.62) the $-\frac{4}{5}$***-law*** is fulfilled, and that

$$S^2(r) = C_{\epsilon_2} r^{2/3 + \mu \frac{1}{9}}\ , \tag{2.63}$$

$$S^3(r) = C_{\epsilon_3} r\ , \tag{2.64}$$

$$S^6(r) = C_{\epsilon_6} r^{2-\mu}\ . \tag{2.65}$$

S^6 is good to estimate the intermittency correction μ.

The structure functions can also be seen as ***spatial*** 2***-point-correlations***,

$$S^n(r) = \langle (u(x+r) - u(x))^n \rangle = \langle u_r^n \rangle = \int_{-\infty}^{\infty} u_r^n p(u_r) du_r\ . \tag{2.66}$$

Thus the structure functions $S^n(r)$ are the general moments of $p(u_r)$.
If the increments u_r are normalized by $\sqrt{< u_r^2 >}$ to $u_r = \sqrt{< u_r^2 >}\, w_r$ we obtain

$$S^n(r) = \int_{-\infty}^{\infty} u_r^n p(u_r) du_r = < u_r^2 >^{n/2} \int w_r^n P(w_r)\, dw_r\ . \tag{2.67}$$

If the integral over w_r is independent of r , this is the case if $P(w_r)$ is the same for all r, then Kolmogorov 1941 is obtained again. The other way round, from the intermittency correction, expressed by μ must have the consequence that $P(w_r)$ is changing its from with r.

2.5.4 Multiplicative cascade after Castaing

We follow here Castaing *et al.* (1990). One studies the ***probability density function*** (PDF) $p(u_r)$ of the velocity increments. One can write $p(u_r)$ as function of the conditioned probability $p(u_r|\epsilon_r)$, as

$$p(u_r) \;=\; \int_0^{+\infty} p(u,\epsilon_r,r)\,d\epsilon_r \;=\; \int_0^{+\infty} p(u_r|\epsilon_r)\,p(\epsilon_r,r)\,d\epsilon_r\,. \tag{2.68}$$

Castaing assumed (all this is experimentally verified, see e.g. Naert *et al.* 1998) that $p(u_r|\epsilon_r)$ is Gaussian distributed:

$$p(u_r|\epsilon_r,r) \;=\; \frac{1}{\sqrt{2\pi}\,s(\epsilon_r)}\,\exp\left(-\frac{u^2}{2\,s^2(\epsilon_r)}\right)\,. \tag{2.69}$$

Next the standard deviation s depends on ϵ_r as

$$s(\epsilon_r) \;\propto\; \epsilon_r^{\alpha}\,. \tag{2.70}$$

As $\ln(s) \propto \alpha \ln(\epsilon_r)$ also s must be log-normal distributed. Thus

$$p(u_r) \;=\; \frac{1}{2\pi\lambda(r)}\int_{-\infty}^{+\infty} \exp\left(-\frac{\ln^2\left(s/s_0(r)\right)}{2\lambda^2(r)}\right)\exp\left(-\frac{u^2}{2s^2}\right)\frac{d\ln s}{s}\,. \tag{2.71}$$

There are two parameters $s_0(r)$ and $\lambda^2(r)$. The first one $s_0(r)$ is the maximum of the distribution of s and determines the variance of $p(u_r)$. The second one $\lambda^2(r)$ is the variance of the log-normal distribution for s. It determines the form of $p(u_r)$ and is thus called the ***form parameter***. In the limit $\lambda^2(r) \to 0$ a Gaussian distribution is obtained. This is the case for $r > L$, as shows the figure 2.15a. On the contrary for small scales r the values λ^2 increase thus a departure from Gauss is seen. By fitting the $p(u_r)$, on can estimate the parameter $\lambda^2(r)$ (figure 2.15b). Thus informations on $\Lambda(r)$ can be obtained, since, as shown in Castaing *et al.* (1990), $\Lambda = 3\lambda$.

2.6 Data analysis - Wind data

As can be deduced from the previous Section, ***wind turbulence*** leads to ***power fluctuations***, ***fatigue*** and ***extreme loads*** on wind Turbines. Therefore a thorough description and characterization of wind turbulence is crucial for a reliable design and efficient operation of wind Turbines. On the other hand, due to the presence of many phenomena at many different scales in the atmosphere, a comprehensive and direct analytical description of turbulence remains a challenge.

A way to handle this complexity is to describe the phenomena in terms of statistical quantities. However which statistical information is necessary and sufficient for a certain purpose is not a trivial question. A useful characterization should enable comparison of wind situations between different sites and allow for the adequate selection of a wind Turbine class and layout. IEC (2005) defines a procedure to achieve these requirements based on 10 minutes mean values of the wind speed and respective deviations from this mean. Definitions are given for wind Turbine classes, wind situations and probability estimations of certain extreme events. While necessity and usefulness of standards are beyond doubt, in the recent years growing demand for a more comprehensive and more detailed characterization has become evident. This is an active field of research, for example Bierbooms (2009) worked on the extreme operative gust and how to incorporate this in the generation of synthetic wind time series. Also recently in Larsen & Hansen (2008) calibrations

of the extreme cases with return periods of 50 and 1 years in the standard IEC 61400 were presented. Common to the cited works is the explicit role of the particular wind Turbine under consideration.

Here we try to review and clarify which kind and which amount of information is grasped by different levels of statistical descriptions of wind turbulence. Our aim is not to give a detailed review of every different statistical approach, but to provide a general and systematic framework which allows to classify and distinguish these approaches and their reach. This is particular important for a systematic review of the statistical methods present in the mentioned IEC norm, while recognizing the value and power of the methods therein but also in the realization of their limitations and where they stand within a systematic and consistent statistical description of wind turbulence. Such statistical framework is also important when defining benchmarks for synthethic and simulated turbulent wind fields.

The structure of this section is such, that with every subsection more statistical information for the description of the wind turbulence is taken into account, therefore enhancing the classification of wind turbulence gradually from low order one-point statistics to n^{th}-order two-point statistics, look at Tab. 2.1 page 80. Note also that we differentiate between statistics taken from some data point x_i or taken from pairs of data like x_i, x_{i+n}, the latter is denoted as two-point statistics. Based on examples we show the suitability and completeness of every step of the statistical description. Thus we start with one-point statistics proceed with two-point statistics and finally comment on n-point statistics. Each time lower and higher order moments of the statistics are discussed.

For illustration purposes we use wind data of the research platform FINO 1[21]. Trying to avoid wake effects caused by the measurement tower we use data from the top anemometer at 100 m height. The wind speed is recorded with a sampling rate of 1 Hz, for our examples we used one month of data namely January 2006. For simplicity we work just with the horizontal magnitude $u(t)$ of the wind velocity vector. It is straight forward to extend our scheme to different directional components of the wind vector, as well as to spatial instead of temporal separations. To demonstrate the importance of this method of data analysis, we compare the results from measured data with a synthetic time series, which was obtained from a common package used by the wind energy community called TurbSim[22].

2.6.1 One-point statistics

One-point statistics up to second order

Given a wind time series, a common approach in the context of wind energy is to define ***turbulent fluctuations*** $u'(t)$ superimposed over a ***mean wind speed*** (Burton *et al.* 2001),

$$u(t) = \langle u \rangle_T + u'(t)_{\langle u \rangle_T}. \tag{2.72}$$

Here $\langle \cdot \rangle$ denotes time average and T the particular time span taken for calculating the average. In general $\langle x^n \rangle$ is the n^{th} moment of x. These fluctuations around the mean value have themselves a mean value of zero, $\langle u'(t)_{\langle u \rangle_T} \rangle = 0$. figure 2.17 (a) shows a typical wind time series and its 10 minutes mean values. figure 2.17 (b) shows the corresponding fluctuations $u'(t)$. Note that in our notation the index in $u'(t)_{\langle u \rangle_T}$ makes explicit the fact that these fluctuations are always defined through $\langle u \rangle_T$. In our case we take a simple average over the time window T, filtering or more elaborated detrending methods would change the actual value of u' and there statistical properties.

[21] www.fino1.de/en .

[22] https://nwtc.nrel.gov/TurbSim .

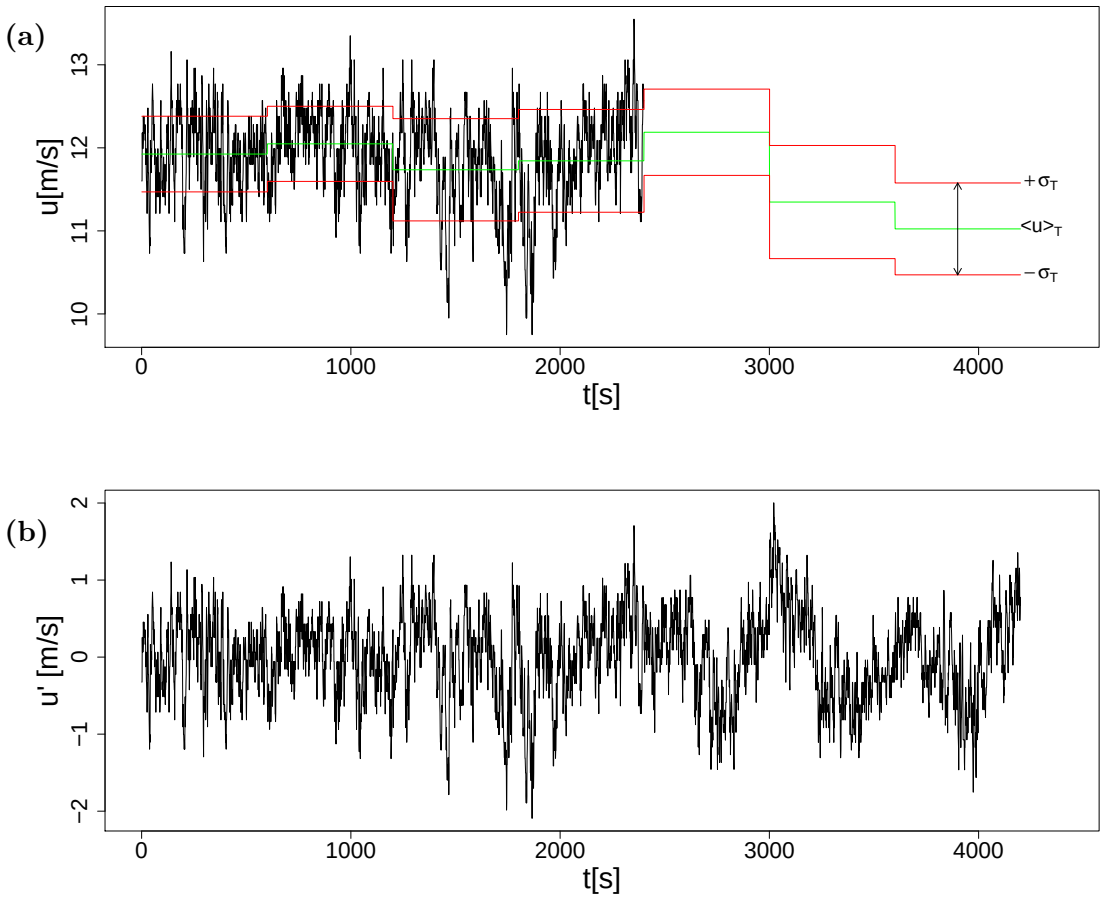

Fig. 2.17 : (a) Time series of the horizontal wind speed $u(t)$ measured on the research platform FINO I at 100 m height, with 10-minute mean values and standard deviations. (b) Fluctuations $u'(t)$ defined over 10-minute mean values.

One issue with this definition is the need of an averaging period T over which to define the mean wind speed used as reference for the fluctuations. Due to the strong non-stationarity of the wind, this task is not trivial, but 10 minutes spans are the usual practice. This particular time span is often motivated by the so-called 'spectral gap', cf. Stull (1994). It is assumed that there is a clear cut between mesoscale variations (large-scale meteorological patterns) and high frequency fluctuations. Another argumentation for an averaging time scale T could be attributed to the response dynamics of a wind Turbine. Wind Turbines can follow adiabatically changes in the wind acting on sufficiently large time scales. In this way one might propose a wind Turbine based time span T, which would change from wind Turbine to wind Turbine, depending on size and control system.

For simplicity, from now on we will just write $u'(t)$ for $u'(t)_{\langle u\rangle_T}$, and we will also adopt the usual $T = 10$ min. Our following arguments could be applied to any other characteristic or desired value of T.

In addition to the 10 minutes mean values $\langle u\rangle_{10min}$, an estimation of turbulence strength in

the wind during the time span T is the ***turbulence intensity***

$$\boxed{I = \frac{\sqrt{\langle u'^2 \rangle}_T}{\langle u \rangle_T} = \frac{\sigma_T}{\langle u \rangle_{10min}}}. \tag{2.73}$$

Here σ_T is the standard deviation of the fluctuations departing from the mean wind speed in the considered period of time. The turbulence intensity is a crucial parameter by which, e.g., certification and site assessment procedures are defined, and it is indeed a design constraint IEC (2005). Details on how this parameter changes with height, mean wind speed or surface roughness can be found for example in IEC (2005); Hansen & Larsen (2005) and for offshore data in Türk & Emeis (2007). As important as it is, the statistical information of this parameter is restricted to the standard deviation or the second moment of the fluctuations. Formally the statistics contained in the turbulence intensity are one-point statistics of second order. As we will introduce in the next section, higher order moments of the fluctuations, $\langle u'^n \rangle$ with $n > 2$, are necessary for the description of extreme values of u'.

Higher order one-point statistics

In the previous subsection we discussed first-order ($\langle u \rangle_{10min}$) and second-order ($\sigma_T$) one point statistics. In general higher order moments are also significant, and this information is contained in the form of the probability density function (PDF) of fluctuations, $p(u')$. Generally, for a PDF there are as characterising quantities, its mean value, its width or standard deviation and its form, which may be Gaussian or not and which is best shown in a normalized presentation, $p(u'/\sigma)$. Note, $\langle x^n \rangle = \int x^n p(x) dx$, the complete set of moments is contained in the knowledge of the PDF of a statistical quantity. The only case where the first two moments will give a complete description of one-point statistics, is the case where the PDF of the fluctuations follows a Gaussian distribution, since only this distribution is completely defined by its first two moments.

The question whether wind fluctuations follow a Gaussian distribution or not becomes relevant in order to understand how much parameters like the turbulence intensity I are needed to characterize such fluctuations. To consider the Gaussianity of the fluctuations, figure 2.18 presents the PDFs for different sets of $u'(t)$. figure 2.18a shows the PDFs for some arbitrary single 10-minutes intervals. Which can be understood as $p(u'(t)|\sigma_T^i)$, where σ_T^i denotes the standard deviation of the i^{th} 10-minutes interval. Within such intervals fluctuations seem to follow the Gaussian distribution, and the pdfs can be characterized by the corresponding σ_T^i. In contrast to this in figure 2.18b the complete set (20 days) of $u'(t)$ is shown. We see that in this case the probabilities of large values of fluctuations are clearly underestimated by a Gaussian distribution. Note the logarithmic y-axis and note that the observed $10\sigma_\infty$ events are underestimated by the Gaussian distribution by a factor of 10^8 i.e. this probability difference translates into the event ocurring once every week or once every $2\ 10^6$ years.

Finally figure 2.18c shows again the PDF of the total data set, but here in each i^{th} 10-minutes interval the fluctuations $u'(t)$ were normalized by the corresponding σ_T^i standard deviation of the respective i^{th} 10-minutes interval. For these locally rescaled fluctuations, the resulting PDF $p(u'(t)/\sigma_T^i)$ is well described by a Gaussian distribution within $\pm 5\sigma$. Small deviations may be seen for the largest values, but the significance here is rather questionable. Nevertheless comparing figure 2.18b and c we clearly see that for $p(u'(t)/\sigma_T^i)$ a Gaussian distribution is a very good approximation.

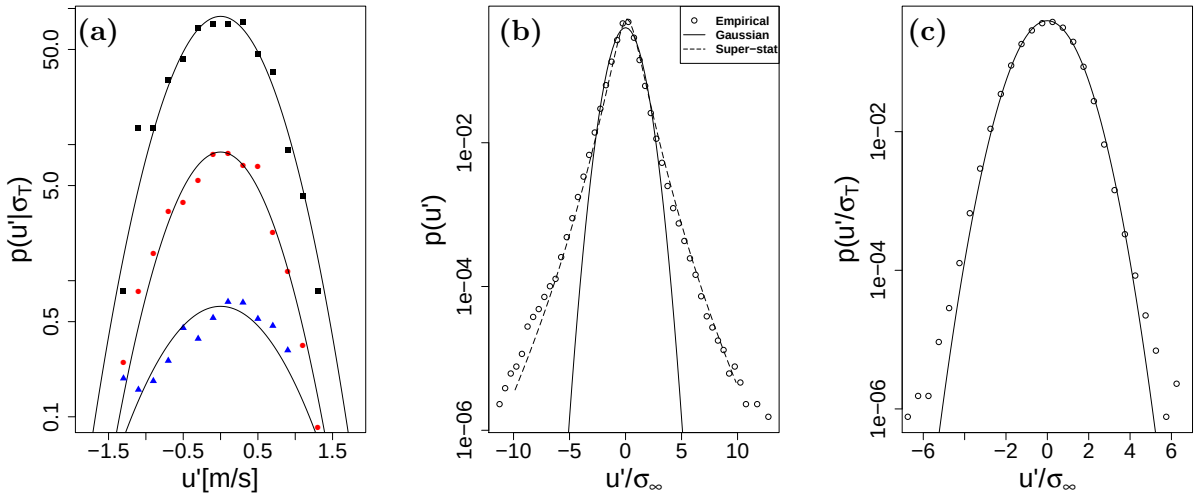

Fig. 2.18 : (a) $p(u'(t)|\sigma_T)$ for three arbitrary ten-minute intervals. From top to bottom the values of the respective σ_T are 0.456, 0.452 and 0.616 m/s, the flatness $\langle u'^4\rangle/\sigma_T^4$ are 2.82, 2.82, and 2.89 ± 0.79. (b) PDF of $u'(t)$ considering all the data set. The flatness is 6.30 ± 0.11, and $\langle (ln\sigma'_T - \overline{ln\sigma'_T})^2\rangle = 0.2037$. (c) PDF of $u'(t)/\sigma_T$ where σ_T is the standard deviation in each interval of length T (see also discussion with the text). Flatness is 2.98 ± 0.01. Symbols represent wind data and solid lines Gaussian distributions in all figures.

These findings support the hypothesis that $u'(t)$ is Gaussian distributed within ten minute intervals, but with different standard deviations σ_T for each interval. It is known that the standard deviations σ_T of $u'(t)$ for single ten-minute intervals are closely log-normal distributed and the parameters of the log-normal distributions depend on the mean wind speed (Hansen & Larsen 2005). Thus if fluctuations $u'(t)$ are considered without normalization, the superposition of the different Gaussians distributions with different σ_T leads to the intermittent distribution in figure 2.18b.

For a more quantitative evaluation of deviations from Gaussianity, it is common to calculate the third and fourth moments of the fluctuations normalized by the standard deviation, which are called ***skewness*** *Skew* and ***flatness*** F, respectively. For a general signal $x(t)$ the definitions are

$$Skew(x) = \frac{\langle (x-\bar{x})^3\rangle}{\sigma_x^3} \tag{2.74}$$

$$F(x) = \frac{\langle (x-\bar{x})^4\rangle}{\sigma_x^4}\,. \tag{2.75}$$

A Gaussian distribution has a skewness value of zero and a flatness of three. For all the cases in figure 2.18, the values of the skewness do not differ significantly from zero, confirming the symmetry of the distributions. The values of the flatness are 2.82, 2.82, and 2.89 ± 0.79 from top to bottom in figure 2.18a and thus do not contradict the Gaussian distribution for arbitrary single ten-minute intervals, while the large deviations from Gaussianity in figure 2.18b result in a flatness of 6.30 ± 0.11. For the PDF of all the rescaled ten-minute intervals in figure 2.18c the flatness of 2.98 ± 0.01 is again surprisingly close to the Gaussian value.

From the discussion above it follows that an assumption of Gaussianity for $u'(t)$ holds only for single ten-minute intervals, and caution should be taken when estimating the probability of extreme values of u', where the actual values of u' and not the rescaled vaules u/σ_T are relevant. Here general higher moments than $\langle u'^2\rangle$, or related quantities like the flatness (see eq. 2.75), will be needed for a correct description of $p(u')$. This non-Gaussianity of the extreme excursions from

the mean wind speed has been already noted in, e.g., Panofsky & Dutton (1984). More recently in Larsen & Hansen (2006) an asymptotic expression for describing the distribution of such extreme events was presented.

Based on our above findings that $p(u'(t)|\sigma_T)$ can be approximated by a Gaussian, we can apply a superposition approach similar to Castaing *et al.* (1990),

$$p(u') = \int p(u'|\sigma'_T) \cdot p(\sigma'_T) d\sigma'_T. \tag{2.76}$$

Here $p(u'|\sigma'_T)$ represents a Gaussian distribution and $p(\sigma'_T)$ a log-normal distribution. The key parameter is $\langle(\ln \sigma'_T - \overline{\ln \sigma'_T})^2\rangle$ which gives directly a measurement of the intermittency of $p(u')$ and is therefore actually related to the flatness of the distribution. figure 2.18b shows that our model based on equation (2.76) is able to describe the PDF $p(u')$ properly even in the tails.

Up to this point we have been discussing stepwise one-point statistics of first order $\langle u \rangle$, of second order in σ_T, and of higher orders summarized in the PDF of u'. However even a complete knowledge of $p(u')$, or respectively of all the moments $\langle u'^n \rangle$, is not unique in the sense that many different time series can share these statistics. To make this point more clear we refer to figure 2.19, which shows three time series which have Gaussian distributions $p(u')$. Those time series share the same one-point statistics, thus the same value of standard deviation and will give the same value of turbulence intensity once added to the same mean wind speed. Clearly the nature of these time series is different.

This is not surprising since the PDF of $u'(t)$ gives no information regarding which path the process follows in order to achieve the observed distribution. In order to distinguish more features of the time series a proper correlation analysis is necessary. The next section will therefore deal with the characterization of correlations by two-point statistics.

2.6.2 Two-point statistics

Two-point statistics up to second order

As seen previously and in figure 2.19 it is in general necessary to obtain knowledge on the correlations between two points in the time series of wind fluctuations. The basic statistical tool for this purpose is the autocorrelation function

$$R_{u'u'}(\tau) \;=\; \frac{1}{\sigma^2_{u'}} \left\langle u'(t+\tau)\, u'(t) \right\rangle \tag{2.77}$$

which quantifies the correlation of two data points separated by the time lag τ. The Wiener-Khintchine theorem (Press *et al.* 1996) relates it to the power spectral density $S(f)$ via a Fourier Transformation $\mathcal{F}$, thus both functions contain the same information:

$$S(f) \overset{\mathcal{F}}{\Longleftrightarrow} R_{u'u'}(\tau) \quad \text{with} \quad \sigma^2_{u'} \;=\; \int S(f) \mathrm{d}f\,. \tag{2.78}$$

These are second order two-point statistics and give information on the intensity or amplitude with which different frequencies contribute to the fluctuations. This statistical tool already enables us to distinguish between the signals shown in figure 2.19 despite that they share the same one-point statistics. As we can see in figure 2.19 in the case of the random signal our description would be complete with the turbulence intensity because the PDFs of the fluctuations are Gaussian and the fluctuations are completely uncorrelated, i.e.,

$$R_{u'u'}(\tau) \;=\; \delta(\tau). \tag{2.79}$$

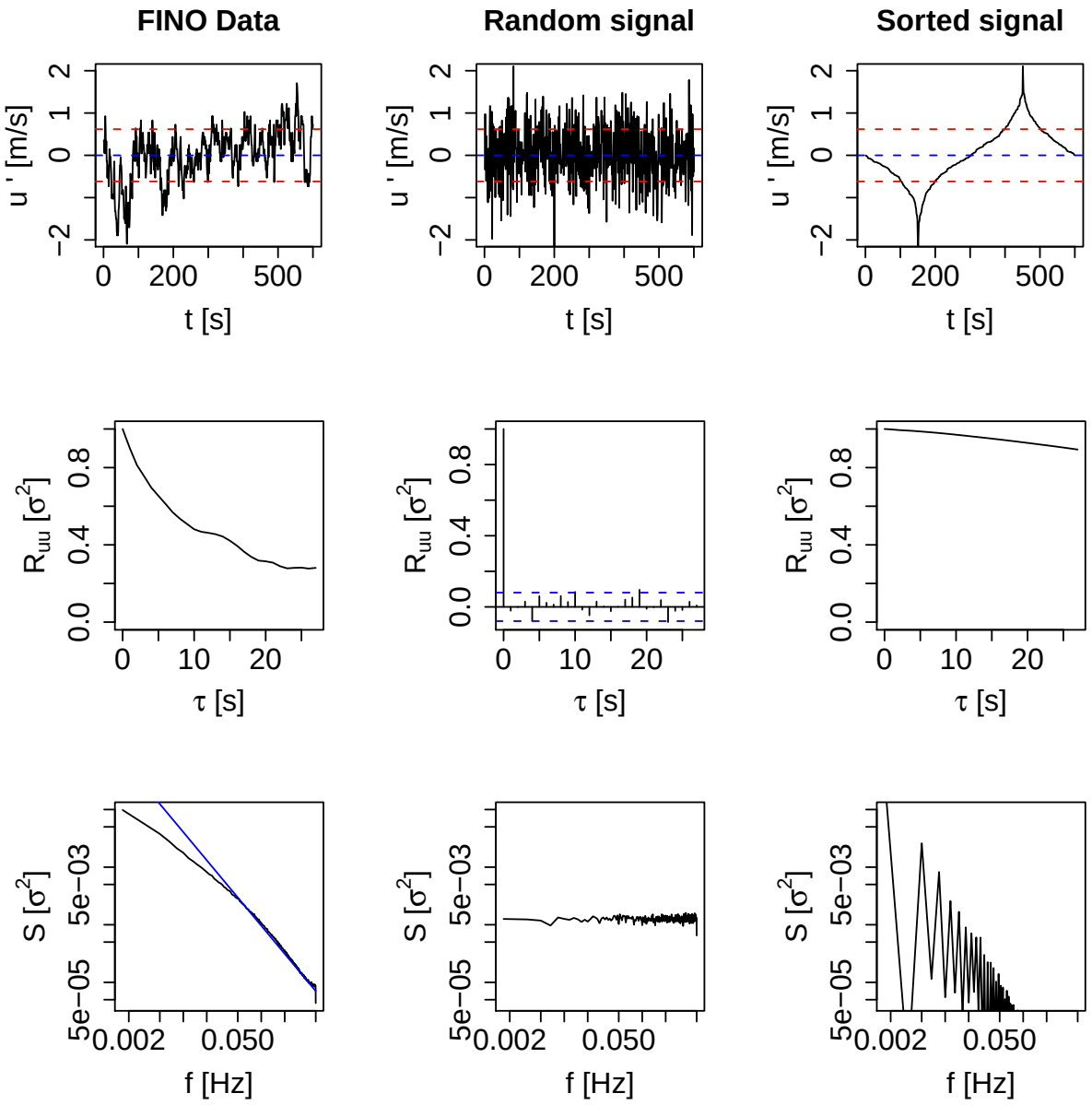

Fig. 2.19 : Three different time series, their autocorrelation functions $R_{u'u'}(\tau) = \langle u'(t+\tau)u'(t)\rangle$, and power spectral densities $S(f)$. From left to right, atmospheric fluctuations u' (measured at FINO 1), a random Gaussian distributed time series, and an ordered time series constructed from the random series. All three share the same standard deviation σ_T, and closely the same $p(u')$. In the top column the dashed lines correspond to one standard deviation. The straight line shown with the FINO 1 power spectral density corresponds to $S(f) \propto f^{-5/3}$. From the figure it is clear how even a complete knowledge of one-point statistics is not sufficient in order to characterize wind turbulence.

However as we know this special case would be very difficult to find in the atmosphere where many different interactions occur over many scales Stull (1994). Instead as can be seen again in figure 2.19 the atmospheric turbulence exhibits a lot of structure regarding the power spectrum. In the so-called inertial range usually a Kolmogorov similarity theory is adopted in order to explain the spectrum over a range of frequencies ($S(f) \propto f^{-5/3}$). In practice either the Kaimal or the von Karman spectra are used not only for the description of atmospheric turbulence but also in the generation of synthetic wind fields.

Similar to the one-point statistics up to second-order summarized in the turbulence intensity I, the autocorrelation function and the power spectral density are important and widely used statistical quantities. Nevertheless, in the general case higher order two-point statistics are indispensable. In principle we should ask ourselves for higher correlations of the form

$$R_{u'^n u'^m}(\tau) \;=\; \frac{1}{\sigma_{u'}^{n+m}} \left\langle u'(t+\tau)^n \; u'(t)^m \right\rangle . \tag{2.80}$$

However, it is more general and even practical to work with the statistics of wind speed differences. As we will see in the next subsection these are also two-point statistics and their moments contain the arbitrary order two-point correlations defined in eq. (2.80). These will be discussed in the next subsection in terms of wind speed differences.

Higher-order two-point statistics

To investigate more generally two-point statistics and higher order correlations in wind turbulence, let us now consider wind speed differences over a specific time lag τ,

$$u_\tau(t) = u(t+\tau) - u(t) = u'(t+\tau) - u'(t) , \tag{2.81}$$

which we will refer to as ***wind speed increments*** in the following. Wind speed increments statistics are clearly by definition two-point statistics, and the necessity of selecting a time span for calculating the mean wind speed is avoided. Instead the increments are defined over a scale τ, and the nature of wind speed variations can be studied against the evolution of this scale. The increment's second moment is directly connected to $R_{uu}(\tau)$ by the simple calculation

$$\langle u_\tau(t)^2 \rangle \;=\; 2\langle u(t)^2 \rangle - 2\langle u(t)\, u(t+\tau) \rangle \;=\; 2\langle u(t)^2 \rangle \left(1 - R_{uu}(\tau)\right) , \tag{2.82}$$

where it is assumed that the time series is long enough to ensure $\langle u(t)^2 \rangle = \langle u(t+\tau)^2 \rangle$ within the desired precision. Note that these considerations apply for the wind speeds $u(t)$ as well as for their fluctuations $u'(t)$, at least inside a ten minute interval, see eq. (2.81). Thus, from the power spectral density and autocorrelation function we obtain the variances or the second moment of the wind speed increments as a function of τ. It is straight forward to see that higher moments of wind speed increments, $\langle u_\tau(t)^n \rangle$ with $n > 2$, are related to higher order correlations, compare with eq. (2.80).

In figure 2.20a we show PDFs of wind speed increments for different time scales τ, together with Gaussian PDFs which share the same standard deviation. Typically PDFs of atmospheric wind speed increments are non-Gaussian for a wide range of scales and have a special 'heavy-tailed' shape Böttcher *et al.* (2003). As already noted in section 2.6.1, the Gaussian distribution is the only one completely determined by the first two moments. Therefore it becomes clear that for wind speed increments the knowledge of higher-order moments than the second is necessary for

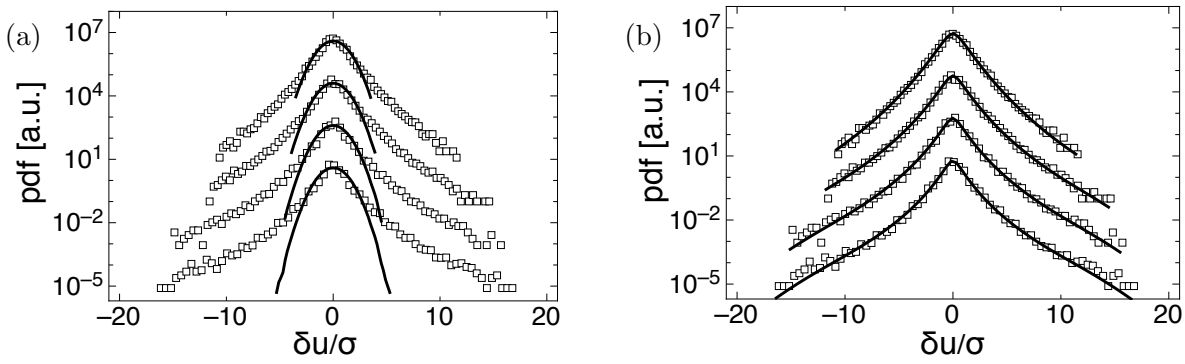

Fig. 2.20 : PDFs of wind speed increments, $u_\tau(t)$, displayed by symbols and shifted vertically. Scales from bottom to top are $\tau = 2$, 4, 10, and 60 s. Additionally, the PDFs are modeled by (a) Gaussian distributions with identical standard deviations, and (b) Castaing's formula (2.83) with (2.84) and (2.85).

a proper characterization of their PDFs. This is a well known and heavily discussed phenomena for turbulence (Frisch 1995) and this is analogous to the case presented in subsection 2.6.1, where we found that higher moments of one-point statistics $\langle u'^n \rangle$ were needed in order to describe the corresponding u' PDFs. AgaThe observed tails in the PDFs imply an increased probability of extreme events, as much as several orders of magnitudes, compared to a Gaussian distribution. Therefore these tails have to be properly reflected in the statistical description.

To this end we follow Böttcher *et al.* (2003) and parameterize the PDFs using Castaing's model Castaing *et al.* (1990), which with some minor modifications is also an explicit formula for eq. (2.76),

$$p(u_\tau) = \frac{1}{2\pi\lambda(\tau)} \int_0^\infty \frac{\mathrm{d}\sigma}{\sigma^2} \exp\left[-\frac{u_\tau^2}{2\sigma^2}\right] \exp\left[-\frac{ln^2(\sigma/\sigma_0)}{2\lambda^2(\tau)}\right] . \tag{2.83}$$

In this equation the PDF is considered as a continuous superposition of Gaussian distributions with different standard deviations, which are weighted by a log-normal distribution function. The shape of the resulting PDF is determined by the two parameters $\lambda^2(\tau)$ and σ_0. Here, σ_0 fixes the median of the lognormal function, while $\lambda^2(\tau)$ mainly dictates the shape of the distribution and is accordingly called the shape parameter. $\lambda^2(\tau)$ is zero for Gaussian distributions and for positive values intermittent distributions are achieved. In general for an emprically given PDF, both parameters can be estimated straightforwardly by an optimization procedure based on equation (2.83). Chilla *et al.* (1996) and Beck Beck (2004) showed that for the case when log-normal superstatistics is the right model, then $\lambda^2(\tau)$ can be directly estimated from the flatness. Following such procedure for Eq. (2.83) we obtain

$$\lambda^2(\tau) = \frac{\ln\left(F_{u_\tau}/3\right)}{4} \tag{2.84}$$

where F_{u_τ} is the flatness of the increment PDF at a given scale τ, cf. equation (2.75). Considering the moments of Gaussian and log-normal distributions we obtain for σ_0, also from Beck (2004),

$$\sigma_0^2 = \langle u_\tau^2 \rangle \exp\left[-2\lambda^2(\tau)\right] . \tag{2.85}$$

In figure 2.20b we model the PDFs by Castaing's formula (2.83), using (2.84) and (2.85) for a simplified estimation of $\lambda^2(\tau)$ and σ_0. It can be seen that the measured increment PDFs are well reproduced for all scales τ.

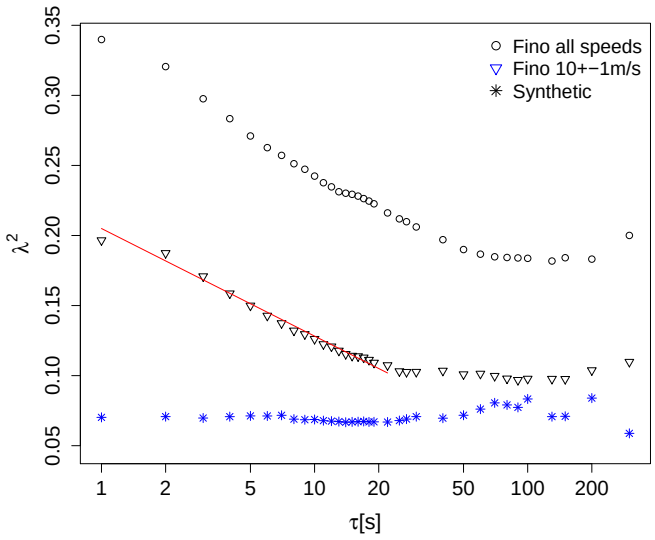

Fig. 2.21 : $\lambda^2(\tau)$ against τ for all wind speeds of the data set ($\circ$), for conditioned wind speeds $(10 \pm 1)\, m/s$ ($\bigtriangledown$), and for a synthetic time series ($*$, see text).

Now, in figure 2.21 we show how the shape parameter behaves against the scale τ for our offshore data. It is important to note the difference between unconditioned and conditioned (by a mean wind speed bin) values of atmospheric data. The behavior of the PDFs or respectively of $\lambda^2(\tau)$ against scale is similar, but the conditioned PDFs show a reduced overall intermittency. The reduction of intermittency for conditioned data sets is due to the fact that part of this intermittency stems from the non-stationarity of the wind.

Regardless of the absolute value of $\lambda(\tau)$ it is important to note that, for both the conditioned and unconditioned sets, there is a clear range of scales τ where $\lambda^2(\tau) \sim \ln \tau$. This logarithmic dependency has a deep meaning in turbulence and is directly related to the intermittency correction of turbulence in the Kolmogorov 1962 theory (Frisch 1995). In particular with

$$\lambda^2(\tau) \approx \lambda_0^2 - \mu \cdot \ln \tau, \tag{2.86}$$

and using eq. (2.83) one gets

$$\langle u_\tau^n \rangle \propto \tau^{\frac{n}{3} - \mu \frac{n(n-3)}{18}}, \tag{2.87}$$

which is the well known multifractal behaviour of turbulence, see the discussion of Kolmogorov 1962 theory at the end of section 2.5.3.

In summary, we have found that higher moments of the increments are neccessary for the proper estimation of the wind speed increments u_τ PDFs, in particular for the correct estimation of extreme events. Fortunately in many cases with $\sigma_{\delta u}(\tau)$ and $F_{\delta u}(\tau)$, we achieve a precise estimation of arbitrary-order two-point statistics of the wind speed. The according wind speed increment PDFs can be modeled following equations (2.83) to (2.85).

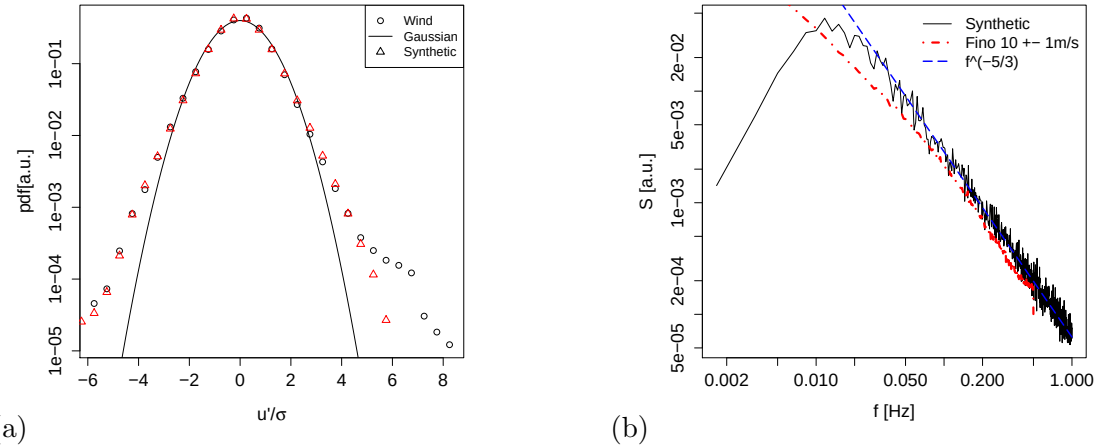

Fig. 2.22 : Basic (a) PDF of $u'(t)$. (b) Normalized FINO and Synthetic spectral densities. A curve following $S \propto f^{-5/3}$ has been added for comparison.

2.6.3 Synthetic time series vs atmospheric turbulence

Previously, we have presented a hierarchical statistical description of atmospheric turbulence. To summarize and contrast this with what is usually standard and used in the wind energy industry, we next apply and compare our statistical scheme to a standard synthetic turbulent wind field and a conditioned FINO data set ($\langle u \rangle_T = 10 \pm 1\, m/s$). We generate with the TurbSim package[22]. 10-min. blocks of synthetic time series with Gaussian pdfs $p(u'|\sigma'_T)$. The 10-minutes blocks are summed up following the distribution $p(\sigma'_T)$ of the conditioned FINO data set (compare with subsection 2.6.1). The resultant synthetic time series reproduce closely $p(u')$ (complete one-point statistics) of the conditioned data as seen in figure 2.22a. We have used TurbSim with the option of a Kaimal power spectrum, therefore the synthetic time series power spectrum follows nicely the law $S(f) \propto f^{-5/3}$, the conditioned atmospheric data follows a similar scaling as seen in figure 2.22b.

On the other hand, when we analyse higher-order two-point statistics, summarized in the behavior of λ^2 against scale τ, we find that the synthetic time series do not at all reproduce the wind PDFs (see figure 2.21). For the synthetic time series the characteristic values of λ^2 remain nearly constant at some low values, i.e. the corresponding PDFs are more Gaussian like.

As discussed in the previous section, the ln dependency of λ^2 against τ, expressed in Eq. (2.86), has an important meaning in terms of the turbulent energy cascade. The failure in reproducing such behaviour by the synthetic time series can not be ignored and represents and important weakness of the models used for the synthetic generation. Moreover, as the response time of different control aspects of a wind Turbine are well within time scales $\tau < 20s$ this part of the scaling plays an important role for the wind Turbine dynamics.

2.6.4 Outlook: n-point statistics

Previously we proposed a comprehensive characterization of increments PDFs by the shape parameter $\lambda(\tau)^2$. These increment PDFs provide information on arbitrary-order two-point correlations. The natural next step in our hierarchical description would be the study of n-scale statistics. The complete stochastic information is contained in the general n-scale joint PDF,

$$p(u_{\tau_1}, u_{\tau_2}, ..., u_{\tau_n}). \tag{2.88}$$

This PDF quantifies the probability that at the same time the wind speed increments $u_{\tau_1}, u_{\tau_2}, ..., u_{\tau_n}$ are observed on the scales $\tau_1, ..., \tau_n$. Note that by the definition of eq. (2.82) and that all increments $u(t,\tau) = u(t+\tau) - u(t)$ share the common point $u(t)$ the n-scale probability (2.88) corresponds to an n-point statistics and would capture the arbitrary n-point correlation.

2.6.5 Conclusions

A statistical characterization of wind turbulence has been presented in a hierarchical and mathematically consistent way. In doing so, we have reviewed the well known one-point statistics up to second order summarized in the turbulence intensity as well as the two-point statistics up to second order reflected in the spectral density. We have shown that in general, in the case of wind speed time series, higher order moments contain relevant statistical information in both one and two-point statistics and can not be ignored. Therefore we propose to use the probability density functions of wind speed fluctuations u' and wind speed increments $u_\tau(t)$ in order to grasp the statistical information of higher moments. In the case of the fluctuations u' we have presented a superposition model which describes the measured PDFs very well. A similar approach is used for the increments statistics. Historically this approach has been already applied in laboratory turbulence by Castaing *et al.* (1990), and for atmospheric turbulence by Böttcher *et al.* (2007). For the characterization of wind speed increments $u_\tau(t)$ it is in many cases possible to characterize these increment PDFs just by the shape parameter $\lambda(\tau)^2$. This parameter can be estimated by the second and fourth moment as shown in eq. (2.84) (this has to be done in a careful way by checking the quality of this approach like in Fig. 2.20). Moreover, after conditioning on certain mean wind speeds, a logarithmic decay of $\lambda(\tau)^2$ has been shown (see figure 2.21) and its meaning on the turbulent cascade has been pointed out. We would like to note that the current practice in wind energy assessment and according regulations (IEC 2005) do not include the characterization of higher-order two point statistics. An easy improvement in the assessment of turbulent conditions would be the systematic estimation of the shape parameter as a function of τ. Of course a question of special practical importance is which scales τ are relevant for WECs. It seems reasonable to expect that these critical scales will depend on every type of machine, however reaction times on the order of seconds can be expected. Thus the scaling behavior of $\lambda(\tau)^2$ presented in Fig. 2.21 results probably relevant for many machines. As shown here and in Mücke *et al.* (2009), common models and simulation packages for generating synthetic wind fields do not reproduce these two-point statistics. As a consequence, quantitative evaluations of possible effects on WECs due to the non-Gaussian behavior of wind speed increments have not been carried out until very recently Mücke *et al.* (2009). An outlook was given on n-point statistics for the characterization of, e.g., gust clustering and the identification of critical wind gusts shapes. Here we have presented an example on the verification of Markov properties of wind speed increment time series. We pointed out how these Markov properties could ease the description of n-point statistics.

Table 2.1 presents a summary of the observed statistical features in wind time series, as well as the statistical parameters we propose to characterize them. Additionally, columns for synthetic time series generated by spectral models and a random time series (see figure 2.19) are presented for comparison purposes.

Type of Statistics	Order	Feature	Random Data	Spectral Models	Wind turbulence	Characterization
1-point	1	Mean speed	•	•	•	$\bar{u} = \langle u(t) \rangle_T$
	2	Turbulence Intensity	•	•	•	$I = \sigma_{u'}/\bar{u}$
	n	Extreme Fluctuations	-	$*$	•	$p(u')$
2-point	2	Distribution of σ_u over f	-	•	•	$S(f) = \mathcal{F}\{R_{uu}(\tau)\}$
	n	Intermittency of $p(\delta u(\tau))$	-	-	•	$\lambda(\tau)$
n-point	n	Arbitrary-order n-point correlations	-	-	•	To be investigated

Tab. 2.1 : Turbulence Characterization Scheme.

2.7 Exercises

These exercises will be done with the program **R** which is free and available for all operating systems. The basic concept is comparable with MatLab, since **R** also works with vectors and matrices, which makes the handling of large data sets quite comfortable. The program can be downloaded from the webpage `www.r-project.org` where also documentation and tutorials can be found.
The **RStudio Desktop** is an extra program that provides an extra user interface. It should be downloaded from `www.rstudio.com/products/rstudio`. It helps e.g. by setting the working directory and also in managing plots generated by the programmed codes.
There are parts in the data analysis that occur more often, e.g. plotting graphs. For this it could be helpful to use and define own `functions()` to save time.

Exercise 2.1 ***Study of aerodynamic forces and losses***

1a Calculate the pressure drop at the rotor plane of a wind turbine operating under ideal Betz conditions (no losses) for an incoming velocity $u_1 = 10$ m/s and an air density $\rho = 1.2$ kg/m^3.

1b Use the result from `1a` to calculate the acting thrust force on a wind turbine with a rotor radius $R = 60$ m and compare it with the expected thrust force at Betz conditions.

1c Use the result from `1b` to calculate the optimal thrust coefficient c_{th} for a machine operating under ideal Betz conditions.

Hint: The thrust force can be seen as a drag force acting on the rotor with u_1 as the reference velocity.

2 Let us assume we have two rotor blades ($R = 40$ m) of three bladed wind turbines with optimal twist distribution along the rotor blade where rotor blade

a) has a constant chord of 4 m

b) has the optimal chord distribution according to the Betz limit.

Calculate for both rotor blades the distribution of the acting lift force along the blade for a constant lift coefficient $c_l = 1.4$, an incoming velocity $u_1 = 8$ m/s, an air density $\rho = 1.2$ kg/m^3 and a tip speed ratio $\lambda = 7$. Additionally we assume operational conditions according to Betz. Plot both distributions on one graph. Discuss the differences.

3 We consider a three bladed wind turbine with optimal chord and twist distribution accounting for the Betz limit. The rotor radius is $R = 50$ m and the turbine is operating at a tip speed ratio of $\lambda = 7$.

3a Calculate and plot the losses (in %) along the rotor blade due to drag forces with

a) a constant lift to drag ratio of $\epsilon = 45$

b) a varying lift to drag ratio according to $\epsilon(r) = 40 + 2.8\sqrt{r}$. Plot both distributions in one graph and compare.

3b Calculate the losses (in %) due to tip vortices.

3c Calculate losses (in %) due to conservation of angular momentum

Compare all three losses - which contributes the most? Is it always like this?

Perform the calculation from `3a` to `3c` with $\lambda = 3$ and $\lambda = 12$.

Exercise 2.2 ***Study of time series of wind speed and wind turbine power***

On the ARCHE page of the module, there are two synchronous time series `wind.txt` and `power.txt`, which contain (normalized) measurements of wind speed u in front of a turbine, and the corresponding (normalized) wind turbine power P at the same time values. The sampling frequency is 1 Hz. The velocity is normalized to its maximum value maxu = 24.35 m/s. The power is normalized to its maximum value of the order of a few MW.

1 Import these files and allocate the time series of the real wind speed u in m/s and of the normalized power P to variables `wind` and `power`. Plot subsets of u and P of about 1000 - 10000 data points in the plane (u, P) to see the evolution and comment.

Hint: You can use a while loop with a counter that increases the limits of the plotted vectors every time the loop runs:

```
i <- 1
inc <- 10000
stopp <- "n"
while (stopp != "y")
{
  plot(...)
  i <- i + inc
  stopp <- readline(prompt = "press y to stop loop: ")
}
```

2 Apply the procedure defined by the IEC norm (IEC 2005), described in section 2.3, to construct the ***IEC power curve*** for this case. Comment.

Hints:

First, to construct the data averaged over time intervals of 10 minutes, you may use a program of the form, where `sf` denotes the sampling frequency of the data and `x` the time series itself:

```
tenminave <- function(x,sf)
{
 windo <- seq(1,length(x),(600*sf))
 xave <- c()
  for (i in 1:(length(windo)-1))
  {
    xave <- c(xave, mean(x[windo[i]:windo[i+1]]))
  }
  return(xave)
}
```

Second, to construct the IEC power curve, you may use a program of the form:

```
breaks <- seq(0, ceiling(max(wind10)), 0.5)
IEC <- rep(0, (length(breaks)-1) )
for (i in 1:length(IEC))
{
  count <- which( (wind10 > breaks[i]) \& (wind10 <= breaks[i+1]) )
  if (length(count) > 0)
 {
    IEC[i] <- mean(power10[count])
  }
}
mids <- seq(breaks[1]+ 0.25, max(breaks)-0.25, 0.5)
points(mids, IEC, col = 2, type = "b", lwd = 3)
```

where `wind10` and `power10` contain already the time averaged information.

Exercise 2.3 ***Study of time series of turbulent flows***

On the web page of the module, there are two time series of turbulent flows, an air into air round free jet - file `Jet.txt` -, and a wind flow - file `Wind.txt` -; moreover on the ARCHE page, there is a time series of a turbulent flow behind a grid `grid.txt`. In all cases we note T the time step between two measurements which is related to the sampling frequency f_s for each measurement. Here we have $f_{s,jet} = 8$ kHz, $f_{s,wind} = 1$ Hz and $f_{s,grid} = 60$ kHz.

1 Import the file `Jet.txt`, `Wind.txt` and `grid.txt` and compute the number N_d of data points.

Hint: The command `data <- as.matrix(read.table(file= " ") )` can be used to read in the data and allocate it to a data-variable to make sure that it can be used directly by other functions.

Solve the following questions for all three time series

2 Plot the first 5000 and the last 5000 data points of the series. Compare and comment. Play also with the number of data points chosen.

3 Determine the mean velocity for the first 5000 and for the last 5000 data points. Compare and comment. Play also with the number of data points chosen.

Hint: Use the `plot() function` and use the options to draw lines instead of points. Since the 1-dimensional data will be stored as a vector, parts of it can be accessed by `data[start:end]`.

4.a Determine the histogram of the data $u(t)$, normalized as a Probability Density Function, and plot it. Indicate the mean value and the standard deviation. Comment.

Hint: Use the function `hist()` and set the options in order to get the Probability Density Function. Plot the results also using the `plot()` function to be able to add vertical lines for the mean and standard deviation using `abline()`.

4.b Plot the PDF of u but in linear - log scales. Comment on the character, Gaussian or non Gaussian, of $u(t)$.

5 Determine the total mean value $\langle u(t) \rangle$ and its standard deviation $\sigma = \sqrt{\langle u'(t)^2 \rangle}$ with the fluctuating velocity $u'(t) = u(t) - \langle u(t) \rangle$.

6 Estimate the ***turbulence intensity*** I and comment.

7 Determine the histogram of the data $u'(t)$, normalized as a Probability Density Function, and plot it in linear - log scales. Comment on the statistical properties of $u'(t)$.

8 Construct a variable with the ***velocity increments***

$$u_\tau(t) = u(t+\tau) - u(t) \ ,$$

for $\tau = 2T$. Determine the histogram of the data $u_\tau(t)$, normalized as a Probability Density Function, and plot it in linear - log scales.
Comment on the statistical properties of $u_\tau(t)$, also, as compared with the ones of $u'(t)$.

9 What do we expect, if we redo the same analysis as in question 8, but with increasing values of τ ?

Solve the following questions just for the grid time series

10 We admit that the time lapse τ corresponding to the ***integral length scale*** L is $\tau_0 = 1040T$. Compute τ_0 in physical units and the integral length scale L according to the Taylor hypothesis.

11 Estimate the ***Reynolds number*** at scale L, $Re = \sigma L/\nu$.

12 Deduce from this and the theory of Kolmogorov 1941 an estimate of the ***dissipation length scale*** η.

Bibliography

Anahua, E., Barth, S. & Peinke, J. 2008 Markovian power curves for wind turbines. *Wind Energy* **11**, 219–232.

Beck, C. 2004 Superstatistics in hydrodynamic turbulence. *Physica D* **193**, 195–207.

Betz, A. 1927 Die Windmühlen im Lichte neuerer Forschung. *Die Naturwissenschaften* **15**, 46.

Bianchi, F., De Battista, H. & Mantz, R. 2006 *Wind Turbine Control Systems*, 2nd edn. Berlin: Springer.

Bierbooms, W. 2009 Constrained Stochastic Simulation of Wind Gusts for Wind Turbine Design. PhD thesis, Delft University Wind Energy Research Institute.

Böttcher, F., Peinke, J., Kleinhans, D. & Friedrich, R. 2007 Handling systems driven by different noise sources – Implications for power estimations. *Wind Energy*, pp. 179–182. Berlin: Springer.

Böttcher, F., Renner, C., Waldl, H.-P. & Peinke, J. 2003 On the statistics of wind gusts. *Boundary-Layer Meteorology* **108**, 163–173.

Bradshaw, P., Ferriss, D. & Atwell, N. 1967 Calculation of boundary-layer development using the turbulent energy equation. *J. Fluid Mech.* **28**, 593–616.

Burton, T., Sharpe, D., Jenkins, N. & Bossanyi, E. 2001 *Wind Energy Handbook*. New York: Wiley.

Castaing, B., Gagne, Y. & Hopfinger, E. 1990 Velocity probability density functions of high Reynolds number turbulence. *Physica D* **46**, 177–200.

Chien, K.-Y. 1982 Predictions of channel and boundary-layer flows with a low-Reynolds-number turbulence model. *AIAA Journal* **20**, 33–38.

Chilla, F., Peinke, J. & Castaing, B. 1996 Multiplicative Process in Turbulent Velocity Statistics: A Simplified Analysis. *J. Phys. II France* **6**, 455–460.

Einstein, A. 1905 Über die von der molekularkinetischen der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. Phys.* pp. 549–560.

Eitel-Amor, G., Örlü, R. & Schlatter, P. 2014 Simulation and validation of a spatially evolving turbulent boundary layer up to $Re_\theta = 8300$. *Int. J. Heat Fluid Flow* **47**, 57–69.

El Khoury, G. K., Schlatter, P., Noorani, A., Fischer, P. F., Brethouwer, G. & Johansson, A. V. 2013 Direct numerical simulation of Turbulent Pipe Flow at moderately high Reynolds numbers. *Flow, Turb. Comb.* **91**, 475–495.

Farias, R. L. S., Ramos, R. O. & da Silva, L. A. 2009 Stochastic Langevin equations: Markovian and non-Markovian dynamics. *Phys. Rev. E* **80**, 031143 (14p).

Fox, R. F. 1977 Analysis of nonstationary, Gaussian and non-Gaussian, generalized Langevin equations using methods of multiplicative stochastic processes. *J. Stat. Phys.* pp. 259–279.

Friedrich, R. & Peinke, J. 2009 Fluid Dynamics, Turbulence. *Encyclopedia of Complexity and Systems Science* (ed. R. A. Meyers), pp. 3641–3661. Springer New York.

Friedrich, R., Peinke, J., Sahimi, M. & Tabar, M. R. R. 2011 Approaching complexity by stochastic methods: From biological systems to turbulence. *Phys. Rep.* **506**, 87 – 162.

Frisch, U. 1995 *Turbulence. The legacy of Kolmogorov*. Cambridge University Press.

Gardiner, C. W. 1985 *Handbook of Stochastic Methods*. Springer Verlag.

Gottschall, J. 2009 Modelling the variability of complex systms by means of Langevin processes. PhD thesis, Carl von Ossietzky Universität Oldenburg, Germany.

Gottschall, J. & Peinke, J. 2008 How to improve the estimation of power curves for wind turbines. *Environmental Res. Lett.* **3**, 015005 (7p).

Hanjalić, K. & Launder, B. 2011 *Modelling turbulence in engineering and the environment: second-moment routes to closure*. Cambridge University Press.

Hansen, K. S. & Larsen, G. C. 2005 Characterising Turbulence Intensity for Fatigue Load Analysis of Wind Turbines. *Wind Engineering* **29**, 319–329.

Heinz, S. 2018 On mean flow universality of turbulent wall flows. I. High Reynolds number flow analysis. *J. Turbulence* **19**, 929–958.

Heinz, S. 2019 On mean flow universality of turbulent wall flows. II. Asymptotic flow analysis. *J. Turbulence* **20**, 174–193.

Heinz, S. 2020 A review of hybrid RANS-LES methods for turbulent flows: Concepts and applications. *Progr. Aerospace Sciences* **114**, 100597.

IEC 2005 Wind Turbine Generator Systems, Part 12: Wind turbine power performance testing. International Standard 61400-12-1. International Electrotechnical Commission.

Jenny, M. 2017 *Turbomachines - Énergies hydraulique et éolienne*. Cours de Mines Nancy (2A), http://energie.mines-nancy.univ-lorraine.fr/fluides/turbo2a.pdf.

Kolmogorov, A. N. 1942 Equations of turbulent motion in an incompressible fluid. *Selected works of A. N. Kolmogorov*, pp. 328–330. Springer 1991, ed. V. M. Tikhomirov, translated from the Russian by V. M. Volosov.

Larsen, G. & Hansen, K. 2006 The statistical distribution of turbulence driven velocity extremes in the atmospheric boundary layer – Cartwright/Longuet-Higgins revised. *Wind Energy. Proc. Euromech Colloquium 464b, Oldenburg* (ed. J. Peinke, P. Schaumann & S. Barth), pp. 111–114. Springer.

Larsen, G. C. & Hansen, K. S. 2008 Rational Calibration of Four IEC Research 61400-1 Extreme External Conditions. *Wind Energy* **11**, 685–702.

Launder, B. E., Reece, G. J. & Rodi, W. 1975 Progress in the development of a Reynolds-stress turbulence closure. *J. Fluid Mech.* **68**, 537–566.

Launder, B. E. & Spalding, D. B. 1974 The numerical computation of turbulent flows. *Comp. Meth. Appl. Mech. Eng.* **3**, 269–289.

Lee, M. & Moser, R. D. 2015 Direct numerical simulation of turbulent channel flow up to $Re_\tau \approx 5200$. *J. Fluid Mech.* **774**, 395–415.

Lück, S., Renner, C., Peinke, J. & Friedrich, R. 2006 The Markov - Einstein coherence length - a new meaning for the Taylor length in turbulence. *Phys. Lett. A* **359**, 335 – 338.

Menter, F. R. 1994 Two-equation eddy-viscosity turbulence models for engineering applications. *AIAA Journal* **32**, 1598–1605.

Milan, P., Mücke, T., Wächter, M. & Peinke, J. 2010 Two numerical modeling approaches for wind energy converters. *Proc. Comput. Wind Eng. 2010* .

Mücke, T., Kleinhans, D. & Peinke, J. 2009 Influence of small scale wind velocity fluctuations on the aerodynamic alternating loads. *Proceedings of EWEC 2009*.

Naert, A., Castaing, B., Chabaud, B., Hebral, B. & Peinke, J. 1998 Conditional statistics of velocity fluctuations in turbulence. *Physica D* **113**, 73–78.

Panofsky, H. & Dutton, J. A. 1984 *Atmospheric Turbulence - Models and Methods for Engineering Applications.*. Wiley, New York.

Papoulis, A. & Pillai, S. U. 2002 *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education.

Plaut, E. 2021*a* *Mécanique des fluides 2 : ondes, couches limites et turbulence*. Cours de Mines Nancy (2A), http://emmanuelplaut.perso.univ-lorraine.fr/publisp.

Plaut, E. 2021*b* *Transition to turbulence in thermoconvection and aerodynamics*. Cours de Mines Nancy (3A), http://emmanuelplaut.perso.univ-lorraine.fr/publisp.

Plaut, E. & Heinz, S. 2021 Exact eddy-viscosity equation for turbulent wall flows - Implications for CFD models. *AIAA Journal*. https://doi.org/10.2514/1.J060761.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. 1996 *Numerical Recipes in C The Art of Scientific Computing*, 2nd edn. Cambridge University Press.

Rankine, W. 1865 On the mechanical principles of the action of propellers. *Transactions of the Institute of Naval Architects* **6**, 13–39.

Rauh, A. 2008 On the Relevance of Basic Hydrodynamics to Wind Energy Technology. *Nonlinear Phenomena in Complex Systems* **11** (2), 158–163.

Rauh, A. & Peinke, J. 2004 A phenomenological model for the dynamic response of wind turbines to turbulent wind. *J. Wind Eng. Ind. Aerodyn.* **92**, 159–183.

Rauh, A. & Seelert, W. 1984 The Betz optimum efficiency for windmills. *Applied Energy* **17**, 15–23.

Renner, C., Peinke, J. & Friedrich, R. 2001 Experimental indications for Markov properties of small-scale turbulence. *J. Fluid Mech.* **433**, 383–409.

Reynolds, O. 1895 On the dynamical theory of incompressible viscous fluids and the determination of the criterion. *Phil. Trans. Roy. Soc. Lond.* **186**, 123–164.

Richardson, L. F. 1922 *Weather prediction by numerical process*. Cambridge University Press.

Risken, H. 1996 *The Fokker-Planck Equation*. Springer.

Rosen, A. & Sheinman, Y. 1994 The average power output of a wind turbine in turbulent wind. *J. Wind Eng. Ind. Aerodyn.* **51**, 287.

Sokolov, I. 2010 Ito, Stratonovich, Hänggi and all the rest: The thermodynamics of interpretation. *Chemical Physics* **375**, 359 – 363, stochastic processes in Physics and Chemistry (in honor of Peter Hänggi).

Stresing, R., Kleinhans, D., Friedrich, R. & Peinke, J. 2011 Different methods to estimate the Einstein-Markov coherence length in turbulence. *Phys. Rev. E* **83**, 046319 (9p).

Stull, R. B. 1994 *An Introduction to Boundary Layer Meteorology*. Kluwer Academic Publishers.

Townsend, A. 1961 Equilibrium layers and wall turbulence. *J. Fluid Mech.* **11**, 97–120.

Türk, M. & Emeis, S. 2007 The Dependence of Offshore Turbulence Intensity on Wind Speed. *DEWI Magazine* **30**, 10–13.

Twele, J. & Gasch, R. 2005 *Windkraftanlagen*. Teubner B.G. GmbH.

Wilcox, D. C. 1988 Reassessment of the scale-determining equation for advanced turbulence models. *AIAA Journal* **26**, 1299–1310.

Wilcox, D. C. 2006 *Turbulence Modeling for CFD*. DCW Industries.

Yamamoto, Y. & Tsuji, Y. 2018 Numerical evidence of logarithmic regions in channel flow at $Re_\tau = 8000$. *Phys. Rev. Fluids* **3**, 012602, 1–10.

Appendix A

Elements on stochastic theory

This appendix provides the mathematical background for the Langevin-power curve introduced physically in section 2.4.

Complex system can be regarded as systems composed of many microscopic interactions, which lead to some random-like dynamics of macroscopic estimates or macroscopic variables. Let us consider a one-dimensional[1] dynamical system described by the macroscopic variable $x(t)$. The process $x(t)$ is defined as ***stochastic*** because its time evolution is described in a probabilistic sense[2]. Although the sample path $x(t)$ can be continuous, and should be for a purely Markov process, practical applications mostly involve ***discrete signals***. From the continuous process $x(t)$, only N samples x_1, x_2, ..., x_N are known at discrete times $t_1 < t_2 < ... < t_N$. $x(t_i) = x_i$ and $x(t_{i+1}) = x_{i+1}$ are known, yet $x(t)$ for $t_i < t < t_{i+1}$ is unknown. The dynamical system is then described by the discrete samples x_i, so that a complete (statistical) description is given through the joint probability distribution

$$f(x_N, t_N;\ x_{N-1}, t_{N-1};\ ...;\ x_1, t_1)\,, \tag{A.1}$$

where $f(A; B; C)$ is the probability of A and B and C happening. The value of sample x_i at time t_i is stochastic, but its probability to have a given value is fixed. Similarly, one can define the conditional probability p following

$$f(x_N, t_N;\ x_{N-1}, t_{N-1};\ ...;\ x_1, t_1) \ = \ p(x_N, t_N | x_{N-1}, t_{N-1}; ...; x_1, t_1)\ f(x_{N-1}, t_{N-1};\ ...;\ x_1, t_1)\,, \tag{A.2}$$

where $p(A|B; C)$ is the probability of A happening conditioned on (given) B and C happen.

The simplest stochastic process that can be thought of is a purely random process with independent samples x_i (Risken 1996). Independence implies that $p(x_i, t_i|Y) = f(x_i, t_i)$ for any arbitrary condition Y, that is, no matter what the condition Y is, x_i will not depend on it. As a consequence, the joint probability of an independent process is

$$f(x_N, t_N\ ; x_{N-1}, t_{N-1};\ ...;\ x_1, t_1) \ = \ \prod_{i=1}^{N} f(x_i, t_i)\,. \tag{A.3}$$

[1] Only an account of one-dimensional stochastic systems is given here, as it suffices for the problem at hand.

[2] Under some weak conditions (that are not presented here for the sake of brevity), if one would let several realizations $x_a(t)$, $x_b(t)$, ..., $x_z(t)$ of that process x evolve in time, one would see that at a given future time t', the exact values of the process have a random character, i.e., $x_a(t') \neq x_b(t') \neq ... \neq x_z(t')$. Yet the probability to obtain a given value remains fixed, giving for the probability distribution $f(x_a, t') = f(x_b, t') = ... = f(x_z, t')$.

A.1 Markov property

Besides the trivial case of an independent process, the next simplest process is a ***Markov process*** (Risken 1996). For a Markov process, only information about the present state is necessary to describe the next future state, regardless of the past state. This means that the state of system x_i at time t_i depends on x_{i-1} at time t_{i-1}, and the conditional probability can be simplified following

$$p(x_i, t_i | x_{i-1}, t_{i-1};\ ...;\ x_1, t_1) \ = \ p(x_i, t_i | x_{i-1}, t_{i-1})\,. \tag{A.4}$$

Equation (A.2) can be rewritten for a Markov process as

$$f(x_N, t_N;\ x_{N-1}, t_{N-1};\ ...;\ x_1, t_1) \ = \ f(x_1, t_1) \prod_{i=2}^{N} p(x_i, t_i | x_{i-1}, t_{i-1})\,. \tag{A.5}$$

The Markov property is often described as ***memoryless***. One should note that a (one-dimensional) Markov process $x(t)$ cannot describe n-order differential systems with $n > 1$[3]. In some cases, the Markov property can emerge by introducing new variables (which remain to be found) to a non-Markov dynamical system, i.e., by making it higher-dimensional.

The total probability theorem gives (Papoulis & Pillai 2002)

$$f(x_i, t_i) = \int dx_{i-1}\, f(x_i, t_i; x_{i-1}, t_{i-1})\,. \tag{A.6}$$

Similarly for the conditional probability

$$\begin{aligned} p(x_i, t_i | x_{i-2}, t_{i-2}) &= \int dx_{i-1}\, p(x_i, t_i; x_{i-1}, t_{i-1} | x_{i-2}, t_{i-2}) \\ &= \int dx_{i-1}\, p(x_i, t_i | x_{i-1}, t_{i-1}; x_{i-2}, t_{i-2})\ p(x_{i-1}, t_{i-1} | x_{i-2}, t_{i-2})\,. \end{aligned} \tag{A.7}$$

Using the Markov assumption in equation (A.4), equation (A.7) becomes the Chapman-Kolmogorov equation

$$p(x_i, t_i | x_{i-2}, t_{i-2}) = \int dx_{i-1}\, p(x_i, t_i | x_{i-1}, t_{i-1})\ p(x_{i-1}, t_{i-1} | x_{i-2}, t_{i-2})\,. \tag{A.8}$$

For non-Markov processes, the future state does not depend only on the present state, but also on a number of past states, see also Risken (1996); Fox (1977); Farias *et al.* (2009). One can define the Einstein-Markov length τ_{mar} (further referred to as Markov length) as the length of the memory kernel, that is the number of past states that influence the present state. Einstein (1905) presents this coarse-graining as a necessary time interval such that the stochastic forces become independent events . This implies the relation

$$p(x_i, t_i | x_{i-1}, t_{i-1};\ ...;\ x_1, t_1) \ = \ p(x_i, t_i | x_{i-1}, t_{i-1};\ ...;\ x_j, t_j) \tag{A.9}$$

with $t_j = t_i - \tau_{mar}$. Experimental signals usually exhibit a non-vanishing, yet finite Markov length, see e.g. Lück *et al.* (2006); Stresing *et al.* (2011) for turbulence. Various Markov tests exist to search for a Markov length in data sets.

[3] Let us consider the example of a deterministic (a special case of stochastic) process $x(t)$ governed by an arbitrary second-order differential equation $\ddot{x} = F(x, \dot{x})$. $x(t)$ is not a Markov process because the future state $x(t + dt)$ does not depend only on the present state $x(t)$, but also on $\dot{x}(t)$. Knowing $x(t)$ does not suffice to know $x(t+dt)$. However, the two-dimensional process $\{x(t), \dot{x}(t)\}$ is a (two-dimensional) Markov process, so that knowing $\{x(t), \dot{x}(t)\}$ suffices to know $\{x(t + dt), \dot{x}(t + dt)\}$.

A.2 Kramers-Moyal expansion

The stochastic process $x(t)$ evolves probabilistically in time. Because of its partly random nature, it is inappropriate to describe its exact time evolution, which is not reproducible. Instead, a description of the probability $f(x,t)$ to find a value x at time t is relevant. For a Markov process, $f(x,t)$ follows a ***master equation***

$$\frac{\partial f(x,t)}{\partial t} = \int dx' \left[w(x' \to x) f(x',t) - w(x \to x') f(x,t) \right], \tag{A.10}$$

where $w(a \to b)$ is the transition rate from state a to state b. More concretely, the law of total probability implies

$$f(x,t+\tau) = \int dx'\, f(x,t+\tau;x',t) = \int dx'\, p(x,t+\tau|x',t) f(x',t) \tag{A.11}$$

with $\tau \geq 0$.

Conditional moments are defined following Risken (1996),

$$M^{(n)}(x',t,\tau) = \int dx\ (x-x')^n\, p(x,t+\tau|x',t) = \left\langle \left[x(t+\tau)-x(t)\right]^n \middle| x(t)=x' \right\rangle \tag{A.12}$$

where $\langle A|B\rangle$ is defined as the mean value of A given that condition B is fulfilled. One can derive the Kramers-Moyal expansion (see complete derivation in Risken 1996)

$$\frac{\partial f(x,t)}{\partial t} = \sum_{n=1}^{\infty} \left(-\frac{\partial}{\partial x} \right)^n D^{(n)}(x,t) f(x,t) \tag{A.13}$$

with the Kramers-Moyal coefficients defined as

$$\begin{aligned} D^{(n)}(x,t) &= \frac{1}{n!} \lim_{\tau \to 0} \frac{1}{\tau} M^{(n)}(x,t,\tau) \\ &= \frac{1}{n!} \lim_{\tau \to 0} \frac{1}{\tau} \left\langle \left[x(t+\tau)-x(t)\right]^n \middle| x(t)=x \right\rangle \\ D^{(n)}(x,t) &= \frac{1}{n!} \frac{\partial M^{(n)}(x,t,\tau)}{\partial \tau}\bigg|_{\tau=0}, \end{aligned} \tag{A.14}$$

where the third relation owes to $M^{(n)}(x,t,\tau=0)=0$.

The ***Kramers-Moyal expansion*** can be formally written

$$\frac{\partial f(x,t)}{\partial t} = L_{KM}(x,t)\ f(x,t) \tag{A.15}$$

with the ***Kramers-Moyal operator*** defined as

$$L_{KM}(x,t) = \sum_{n=1}^{\infty} \left(-\frac{\partial}{\partial x} \right)^n D^{(n)}(x,t)\,. \tag{A.16}$$

Risken (1996) *assumes*[4] that the Kramers-Moyal expansion describes a Markov process, as the evolution $\partial f(x,t)/\partial t$ of the process at time t depends only on its present state $f(x,t)$, and not on some past states $f(x,t')$ for $t' < t$.

[4]Some criticism of the Kramers-Moyal expansion is formulated in Gardiner (1985), where it is argued that the Kramers-Moyal expansion cannot describe the evolution of some Markov jump processes, but only approximate it.

A.3 Fokker-Planck equation

The Fokker-Planck equation is a special case of the Kramers-Moyal expansion for which $D^{(n)}(x) = 0$ for $n \geq 3$. This relates to the Pawula theorem, which states[5] that the Kramers-Moyal expansion (A.13) either stops after $n = 1$, after $n = 2$, or require an infinity of terms, see Risken (1996). This theorem shows from the generalized Schwartz inequality that $[M^{(2n+m)}]^2 \leq M^{(2n)} M^{(2n+2m)}$ for any set of integers $(n, m \geq 0)$. This implies that $D^{(n>2)} = 0$ if there exists one integer $r > 0$ such that $D^{(2r)} = 0$.

If a Markov process $x(t)$ satisfies the Pawula theorem, the Kramers-Moyal expansion stops after the second term and the probability distribution $f(x,t)$ is described by the ***Fokker-Planck equation***

$$\boxed{\frac{\partial f(x,t)}{\partial t} = \left[-\frac{\partial}{\partial x} D^{(1)}(x,t) + \frac{\partial^2}{\partial x^2} D^{(2)}(x,t) \right] f(x,t) = L_{FP}(x,t)\ f(x,t)} \,. \tag{A.17}$$

The ***Fokker-Planck operator*** reads

$$L_{FP}(x,t) = -\frac{\partial}{\partial x} D^{(1)}(x,t) + \frac{\partial^2}{\partial x^2} D^{(2)}(x,t). \tag{A.18}$$

The stationary solution $f_{st}(x)$ of the Fokker-Planck equation, for time-independent coefficients $D^{(1,2)}(x)$, can be derived from

$$\frac{\partial f_{st}(x)}{\partial t} = \left[-\frac{\partial}{\partial x} D^{(1)}(x) + \frac{\partial^2}{\partial x^2} D^{(2)}(x) \right] f_{st}(x) = 0 \,. \tag{A.19}$$

The solution reads

$$f_{st}(x) = \frac{N}{D^{(2)}(x)} \exp\left(\int^x \frac{D^{(1)}(x')}{D^{(2)}(x')} dx' \right) \tag{A.20}$$

with N a normalization constant such that $\int_{-\infty}^{\infty} f_{st}(x) dx = 1$.

Similarly, the Fokker-Planck equation exists for a conditional probability $p(x,t|x',t')$ (that is the distribution $f(x,t)$ for the initial condition $f(x,t') = \delta(x - x')$) and gives

$$\frac{\partial}{\partial t} p(x,t|x',t') = L_{FP}(x,t)\ p(x,t|x',t') \,, \tag{A.21}$$

which has a unique initial condition $p(x,t|x',t) = \delta(x - x')$.

Let us now consider a process that is stationary in time:

$$f(x_N, t_N; ...; x_1, t_1) = f(x_N, t_N + \tau; ...; x_1, t_1 + \tau)$$

for an arbitrary time shift τ. Then $L_{FP}(x,t) = L_{FP}(x)$ and $D^{(n)}(x,t) = D^{(n)}(x)$, so equation (A.21) has a formal solution

$$p(x, t+\tau|x', t) = e^{\tau\ L_{FP}(x)} \delta(x - x') \tag{A.22}$$

owing to the initial condition $p(x,t|x',t) = \delta(x - x')$.

It is shown in Risken (1996) that equation (A.22) also holds for a time-dependent $L_{FP}(x,t)$ if the time increment τ is sufficiently small so that $D^{(n)}$ can be seen as unchanged coefficients. Based

[5] The Pawula theorem only applies if the conditional probability $p(x, t+\tau|x', t)$ is a non-negative function.

on the definition of the delta function $\delta(x-x')$, equation (A.22) gives the short-time propagator of the Fokker-Planck equation for small τ (for details see Risken 1996):

$$p(x,t+\tau|x',t) \;=\; \frac{1}{\sqrt{4\pi\tau D^{(2)}(x',t)}}\;\exp\left(-\frac{[x-x'-\tau D^{(1)}(x',t)]^2}{4\tau D^{(2)}(x',t)}\right). \tag{A.23}$$

The Fokker-Planck equation (A.17) is a linear partial differential equation that can be solved numerically. Besides the direct method that consists in numerically approximating the differential operators, one can use a path integral method (Risken 1996). Similarly to what is done in quantum mechanics to solve the Schrödinger equation, this method is easy to implement. Given an initial condition $f(x,t_0)$, equation (A.11) gives $f(x,t_0+\tau)$ using $p\big(x,t_0+\tau\big|x,t_0\big)$. This can be iterated $n-$times to calculate $f(x,t_n)$ from $f(x,t_{n-1})$ for time $t_n = t_0+n\tau$. One can formulate this following

$$\begin{aligned} f(x,t_n) = \int dx_{n-1}\int dx_{n-2}...\int dx_0 \\ p\big(x,t_n\big|x_{n-1},t_{n-1}\big)p\big(x_{n-1},t_{n-1}\big|x_{n-2},t_{n-2}\big)...p\big(x_1,t_1\big|x_0,t_0\big)f(x_0,t_0)\,. \end{aligned} \tag{A.24}$$

For a small enough time increment τ, the conditional probability is given by equation (A.23). A similar approach can be used for $p(x,t_0+n\tau|x',t_0)$ using the Chapman-Kolmogorov equation (A.8) and the initial condition $p(x,t_0|x',t_0)=\delta(x-x')$.

A.4 Langevin equation

Based on the historical example of Brownian motion, a ***Langevin equation*** for a ***Langevin process*** $x(t)$ reads

$$\boxed{\frac{dx}{dt} = D^{(1)}(x,t) + \sqrt{D^{(2)}(x,t)}\cdot\Gamma(t)}\,. \tag{A.25}$$

The Kramers-Moyal coefficients are the coefficients defined in equation (A.14).

The time evolution of a sample path $x(t)$ is described by the so-called drift coefficient $D^{(1)}(x,t)$ and diffusion coefficient $D^{(2)}(x,t)$. In parallel, the probability $f(x,t)$ is described by the Fokker-Planck equation (A.17). For a given set of drift and diffusion coefficients, the Fokker-Planck equation gives a unique solution $f(x,t)$. On the contrary, the Langevin equation can generate different sample paths $x_i(t)$ that have different values (due to the randomness of the Langevin noise $\Gamma_i(t)$). Yet the probability that $x_i(t_j)=X$ is the unique solution of the Fokker-Planck equation $f(X,t_j)$.

The Langevin equation can be discretized following

$$\begin{aligned} x(t+dt) &= x(t) + \int_t^{t+dt}\frac{dx}{dt}(t')dt' \\ &= x(t) + \int_t^{t+dt} D^{(1)}(x,t')dt' + \int_t^{t+dt}\sqrt{D^{(2)}(x,t')}\Gamma(t')dt'\,. \end{aligned} \tag{A.26}$$

The integration of $\Gamma(t)$ is not defined mathematically, yet a physical interpretation of the stochastic integral is needed. The definition of stochastic integration in the sense of Itô gives

$$\int_t^{t+dt} g(x,t')\Gamma(t')dt' = g(x,t)\int_t^{t+dt}\Gamma(t')dt'\,. \tag{A.27}$$

For $dt \ll 1$, $D^{(n)}$ can be taken as unchanged coefficients and equation (A.26) becomes

$$x(t+dt) = x(t) + D^{(1)}(x,t)dt + \sqrt{D^{(2)}(x,t)} \int_t^{t+dt} \Gamma(t')dt' \,. \tag{A.28}$$

The Langevin noise $\Gamma(t)$ fluctuates much faster than the stochastic process $x(t)$ and has a correlation length much shorter than dt (its theoretical correlation length is zero). The Langevin noise is related to a Wiener process $W(t)$ in that $\dot{W} = \Gamma(t)$, bringing the Stieltjes integral

$$\int_t^{t+dt} \Gamma(t')dt' = \int_t^{t+dt} dW(t') = \sqrt{dt} \cdot \eta(t) \tag{A.29}$$

with $\eta(t)$ a set of independent, Gaussian-distributed samples following $\langle \eta(t) \rangle = 0$ and $\langle \eta(t)^2 \rangle = 2$. The discrete form of the Langevin equation for a small time increment dt becomes

$$x(t+dt) = x(t) + D^{(1)}(x,t)dt + \sqrt{D^{(2)}(x,t)dt} \cdot \eta(t) \,. \tag{A.30}$$

A sample path $x(t_0 + n\ dt)$ can be generated by iterating $n-$times the integration from an initial condition $x(t_0)$.

Remark concerning the stochastic integrals

In equation (A.27) we used the definition of stochastic integration in the sense of Itô. Another common approach to carry out a stochastic integral is the Stratanovich definition that reads

$$\int_t^{t+dt} g(x,t')\Gamma(t')dt' \;=\; g\Big(\frac{x(t+dt)-x(t)}{2}\ ,\ t+dt/2\Big) \int_t^{t+dt} \Gamma(t')dt' \ .$$

The Stratanovich approach is more intuitive because it considers the value of the function g at the middle point of the integration range. Only the Itô interpretation is used here, as it is easier to implement numerically. Also, the definition of the Kramers-Moyal coefficients is different in the Stratanovich interpretation. Both interpretations are equivalent, as they yield identical probability distributions (Friedrich *et al.* 2011). Sokolov (2010) summarizes various interpretations of the Langevin equation in the presence of non-constant $D^{(2)}$.