7 Statistics

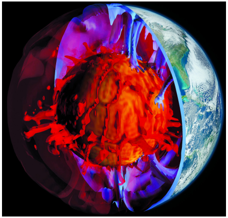
7.1 Introduction

Statistics in the context of CFD

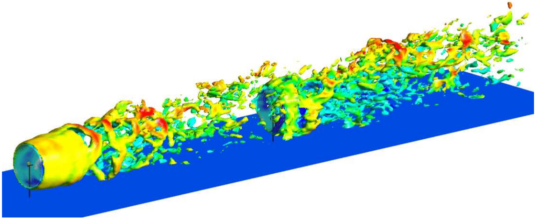
- Fluid flows are often **turbulent** (chaotic)
- A **statistical description** is therefore appropriate

1

2



3



4

1 – NASA, Langley Research Center, public domain
2 – NREL (`www.nrel.gov`)
3 – GEOFLOW, ESA (`www.esa.int`)
4 – NeSI (`www.nesi.org.nz`)

First documented observation of turbulence

Leonardo da Vinci (1452–1519)

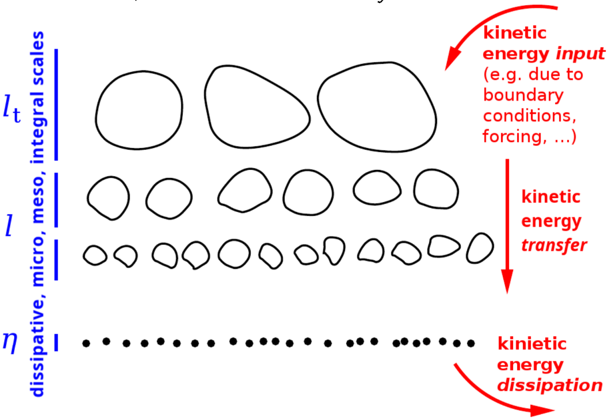
Observe the motion of the surface of the water, which resembles that of hair, which has two motions, of which one is caused by the weight of the hair, the other by the direction of the curls; thus the water has eddying motions, one part of which is due to the principal current, the other to random and reverse motion.[1]

[1] J. L. Lumley, "Some comments on turbulence", *Phys. Fluids A*, 1997, 4:203–211.

A modern view on turbulence phenomenology

Richardson[2] (1922) – description of the energy cascade

Big whirls have little whorls that feed on their velocity, and little whorls have lesser whorls, and so on to viscosity.

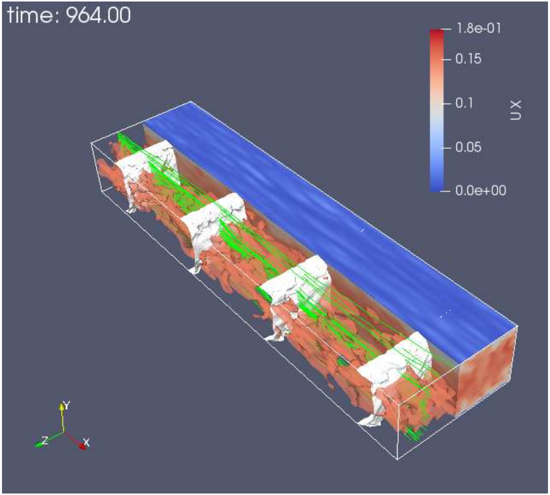


after: J. Fröhlich, TU Dresden

[2] L. F. Richardson, *Weather prediction by numerical process*, Cambridge, 1922.

A canonical case: CFD simulation of turbulent channel flow

Large-eddy simulation (LES) with WALE turbulence model using OpenFOAM



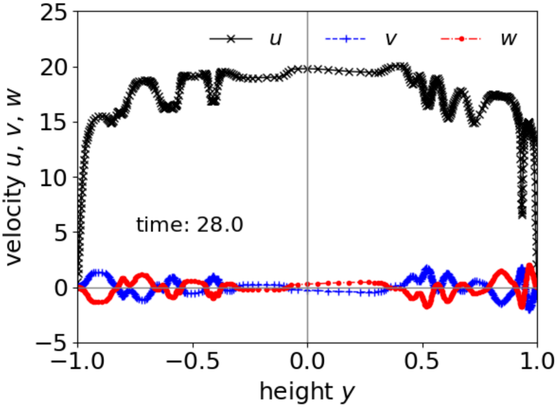
Video in Moodle:

`anim_chan395_OFles3d.mp4`

Statistical turbulence modeling aims to predict the large-scale flow. Statisitcal properties are assumed for the small scales, e.g., based on the Richardson cascade!

Stochastic modeling of turbulent channel flow

One-dimensional turbulence (ODT) simulation



Momentary flow profiles $u(y,t)$, $v(y,t)$, $w(y,t)$

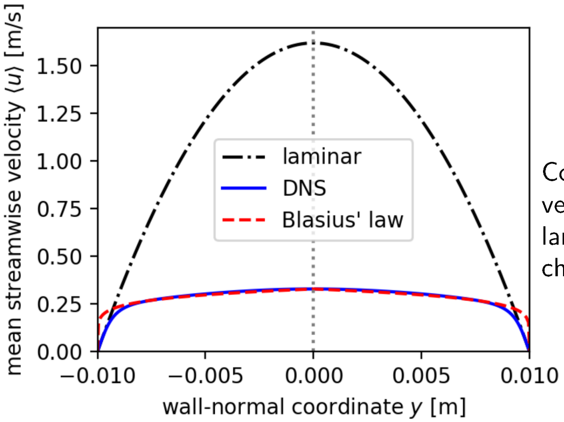
Video in Moodle:

`channel590_odt2.mp4`

Stochastic turbulence modeling aims to reproduce chaotic fluctuations. This requires random numbers sampled from a random process!

Mean velocity profiles for channel flow

- The *fully-developed* turbulent flow is **statistically stationary**
- *Averaging* yields the **mean streamwise velocity** $\langle \boldsymbol{u} \rangle$



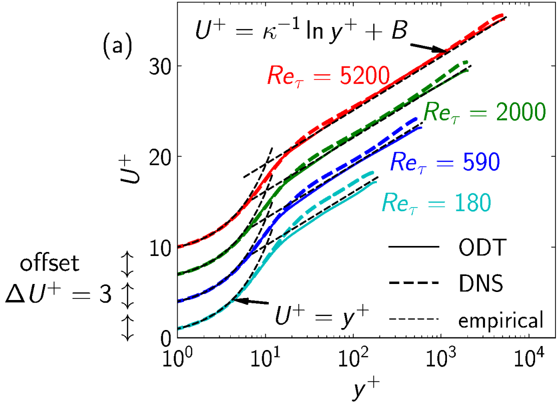
Comparison of mean velocity profiles in laminar and turbulent channel flow

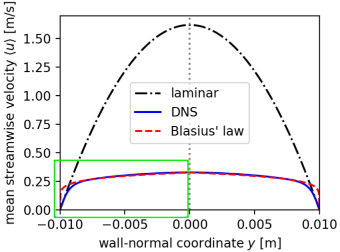
Mean velocity profile across the boundary layer

- Normalized variables (wall units):

$$U^+ = U/u_\tau \qquad y^+ = \frac{u_\tau y}{\nu} \qquad Re_\tau = \frac{u_\tau \delta}{\nu} \qquad \text{(channel: } \delta = h/2\text{)}$$

- Friction velocity scale: $\quad u_\tau = \sqrt{\tau_\mathrm{w}/\rho}$

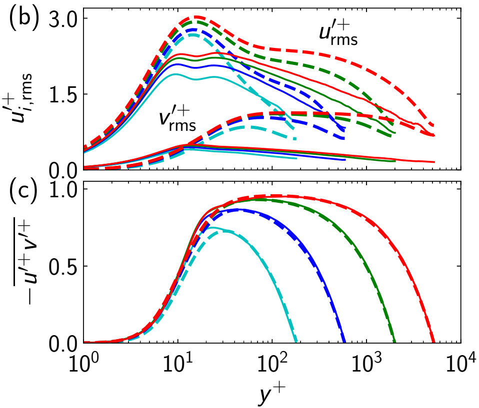




— **ODT:** M. Klein & H. Schmidt (2021) *STAB/DGLR Symposium 2020, NNFM* **151**:47–57
– – **DNS:** M. Lee & R. D. Moser (2015) *J. Fluid Mech.* **774**:395–415

Standard deviation $u'_{i,\mathrm{rms}}$ and cross-correlation $\overline{u'v'}$ of velocity fluctuations across the boundary layer

$$u'_{i,\mathrm{rms}} = \sqrt{\overline{u_i'^{\,2}}}$$



— **ODT:** M. Klein & H. Schmidt (2021) *STAB/DGLR Symposium 2020, NNFM* **151**:47–57

– – **DNS:** M. Lee & R. D. Moser (2015) *J. Fluid Mech.* **774**:395–415

7 Statistics

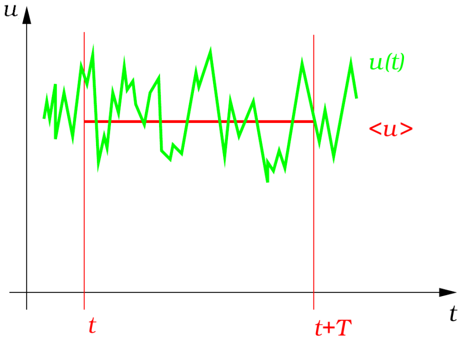
7.2 Definitions

Statistical description of turbulent flows

Separation of flow variables, for example, the velocity vector u_i, into the *mean* $\langle u_i \rangle$ and the *fluctuations* u_i' around the mean

$$\boxed{\boldsymbol{u_i = \langle u_i \rangle + u_i'}}$$

Reynolds decomposition



after: N. Adams, TU Munich

- **Temporal average**

$$\langle u_i \rangle_t = \frac{1}{T} \int_t^{t+T} u_i(\boldsymbol{x}, \hat{t}) \, \mathrm{d}\hat{t}$$

- **Spatial average**

$$\langle u_i \rangle_V = \frac{1}{V} \int_V u_i(\hat{\boldsymbol{x}}, t) \, \mathrm{d}\hat{V}$$

- **Ensemble average**

$$\langle u_i \rangle_N = \frac{1}{N} \sum_{n=1}^{N} u_i^{(n)}(\boldsymbol{x}, t)$$

Mean

- Consider a discrete sample $\{x_1, x_2, x_3, \ldots, x_i, \ldots, x_N\}$ of N values in time, space, or due to an ensembles of different realizations

- The **mean value $\bar{x}$** is given by the arithmetic average

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i$$

- This is the *first statistical moment* of the distribution of the x_i

Variance and standard deviation

- The **fluctuation variance σ^2** is obtained with unbiased estimator

$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$$

N.B.: Why '$N-1$'? Because for only one value, $N=1$, there is no variance. Hence, $\sigma^2 = \frac{0}{0} = \text{n.d.}$

- The **standard deviation of the fluctuations σ** is given by

$$\sigma = \sqrt{\sigma^2}$$

- For large $N \gg 1$, σ^2 can be obtained as difference of the *second* and *first statistical moment* squared

$$\sigma^2 = \left(\frac{1}{N}\sum_{i=1}^{N} x_i^2\right) - \bar{x}^2$$

N.B.:

$$\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\bar{x})^2 = \underbrace{\frac{1}{N-1}}_{\simeq N}\sum_{i=1}^{N}\left(x_i^2 - 2\bar{x}x_i + \bar{x}^2\right) = \frac{1}{N}\sum_{i=1}^{N}x_i^2 \underbrace{-2\bar{x}\frac{1}{N}\sum_{i=1}^{N}x_i}_{-2\bar{x}^2} + \underbrace{\frac{1}{N}\sum_{i=1}^{N}\bar{x}^2}_{\bar{x}^2} = \frac{1}{N}\sum_{i=1}^{N}x_i^2 - \bar{x}^2$$

Higher-order statistical moments

- The equation for the mean of a discrete sample can be generalized to arbitrary statisitcal moments q

$$\overline{x^q} = \frac{1}{N} \sum_{i=1}^{N} x_i^q$$

- $\boldsymbol{q > 2}$ denotes **higher-order statistical moments**

- Correspondingly, higher-order statistical moments of the **fluctuations** are given by the unbiased estimator

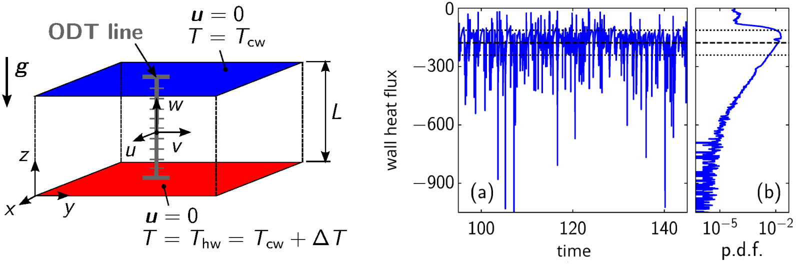
$$\overline{\sigma^q} = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^q$$

- This is used to define additional statisitcal quantities, such as the *skewness* for $q = 3$ or *flatness* (*kurtosis*) for $q = 4$, giving additional information about the distribution (PDF) of the x_i

Histogram and probability density function (PDF) I

- The **histogram** gives the *absolute number* n_k of values x_i that fall within predefined intervals (bins) $[b_k, b_{k+1}]$

- The **probability density function (PDF)** gives the *relative number* $\frac{n_k}{N}$ of values x_i per bin of size $\Delta b_k = b_{k+1} - b_k$, that is, $p_k = \frac{n_k}{N \cdot \Delta b_k}$



Discrete time series (a) and PDF (b) of wall heat flux obtained with a stochastic model for thermal convection. M. Klein, ICTAM 2020+1.
— raw data   – – mean   · · · mean ± standard deviation

Histogram and probability density function (PDF) II

- Algorithmically, n_k is obtained by counting all x_i that fulfill $b_k \le x_i \le b_{k+1}$ for given k. This is repeated for all k.

  $\rightarrow$ Use *nested loops*: k goes to the outer and i to the inner loop

- Note that $k = 1, 2, \ldots, M$ and $i = 1, 2, \ldots, N$, where $N \gg M \gg 1$ for a reasonably populated and resolved histogram that *approximates* the underlying distribution (the 'true' PDF)

- Technically, the PDF is the *normalized histogram* in the limit $N, M \rightarrow \infty$ with $\max_k(\Delta b_k) \rightarrow 0$

Obtaining arbitrary statistical moments with the PDF

- The PDF is the normalized histogram in the infinitesimal limit
- Adopting infinitesimal calculus notation, we have

$$\begin{aligned} \text{discrete variable } x_i &\rightarrow \text{continuous variable } x \\ \text{discrete bin size } \Delta b_k &\rightarrow \text{infinitesimal interval } \mathrm{d}x \\ \text{discrete PDF } p_k &\rightarrow \text{continuous PDF } p(x) \end{aligned}$$

- Any **statistical moment of a continuous variable** is obtained from

$$\overline{x^q} = \int_{-\infty}^{+\infty} x^q \, p(x) \, \mathrm{d}x$$

The PDF contains <u>all</u> statistical information for the considered sample.
This includes the mean <u>and</u> details of fluctuations.

Retrieving low-order moments from the PDF

- Normalization follows from $q = 0$, that is, the integral of the PDF must be equal to one

$$1 = \int_{-\infty}^{+\infty} p(x)\,\mathrm{d}x$$

- The mean is obtained for $q = 1$ as

$$\bar{x} = \int_{-\infty}^{+\infty} x\,p(x)\,\mathrm{d}x$$

- The variance is obtained for $q = 2$ as

$$\begin{aligned}
\sigma^2 &= \int_{-\infty}^{+\infty} (x - \bar{x})^2 \; p(x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} \left(x^2 - 2\bar{x}x + \bar{x}^2\right) \; p(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} x^2\,p(x)\,\mathrm{d}x - 2\bar{x}\int_{-\infty}^{+\infty} x\,p(x)\,\mathrm{d}x + \bar{x}^2\int_{-\infty}^{+\infty} p(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} x^2\,p(x)\,\mathrm{d}x - \bar{x}^2\int_{-\infty}^{+\infty} p(x)\,\mathrm{d}x = \overline{x^2} - \bar{x}^2
\end{aligned}$$

Extension to two and more variables

- Consider time series of two variables $x(t)$ and $y(t)$ that may be represented by discrete samples (x_i, y_i) with $i = 1, 2, \dots, N$
- The PDF will now depend on two variables, it is called the **joint probability density function (JPDF)**

$$P(x, y) \qquad \text{such that} \qquad 1 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(x, y) \, \mathrm{d}x \, \mathrm{d}y$$

- The statistical moments now also encompass **cross-correlations** for $r > 0$ and $s > 0$

$$\overline{x^r y^s} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^r \, y^s \, P(x, y) \, \mathrm{d}x \, \mathrm{d}y$$

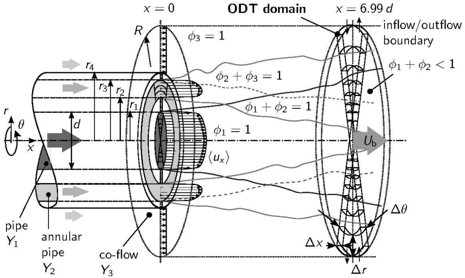
- **Auto-correlations** $\overline{x^r}$ and $\overline{y^s}$ are obtained from the marginal PDFs $p_1(x)$ and $p_2(y)$ in the case of $s = 0$ and $r = 0$, respectively, since

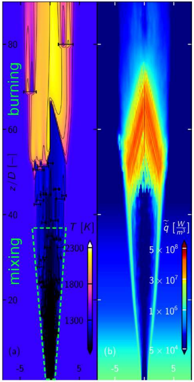
$$p_1(x) = \int_{-\infty}^{+\infty} P(x, y) \, \mathrm{d}y \qquad p_2(y) = \int_{-\infty}^{+\infty} P(x, y) \, \mathrm{d}x$$

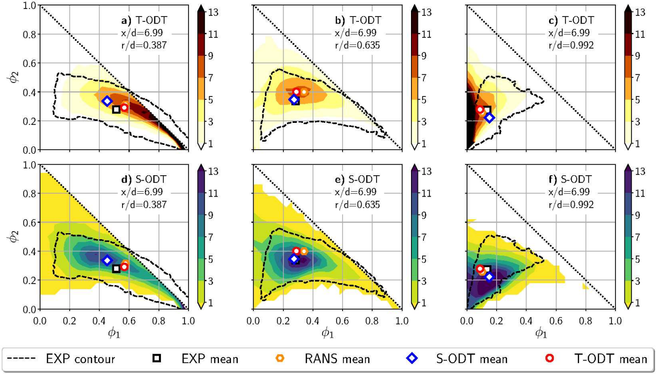
JPDFs of two-scalar mixing in a round jet

M. Klein *et al.*, TSFP12, 2022

T. Starick *et al.*, TSFP11, 2019







7 Statistics

7.3 Hands-on applications

Statisitcal analysis and prediction of dynamical systems

Diagnostics

- Analysis of fluid flows in CFD (DNS, LES) and measurements
- Analysis of the Lorenz system

$\rightarrow$ *Exercise 7*

Prediction

- Formulation of Reynolds-averaged Navier–Stokes (RANS) equations
- Formulation of statistical turbulence models in RANS and LES
- Formulation of PDF transport models for CFD applications

Statisitcal analysis of random processes

- Analysis of random processes (like throw of a dice)
- Analysis of random walks (like Brownian motion, tracer dispersion)
- Quality assurance of random number generators
- ...

$\rightarrow$ *Exercise 8*

Keywords

- Statistics of turbulent fluid flows
- Mean and variance
- Higher-order moments
- Probability density function (PDF)