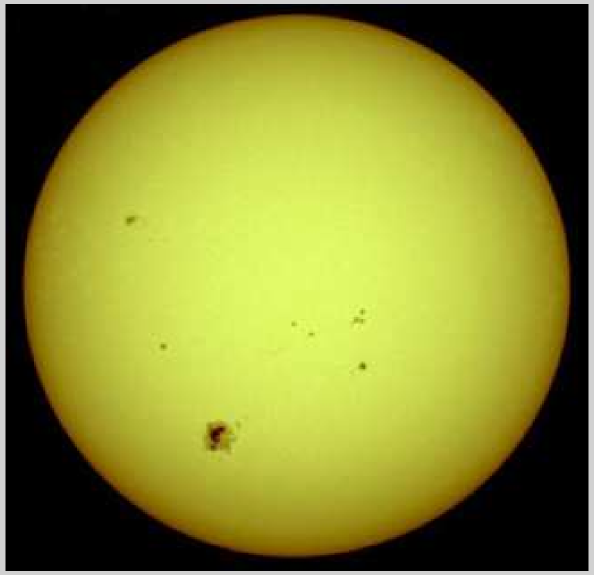
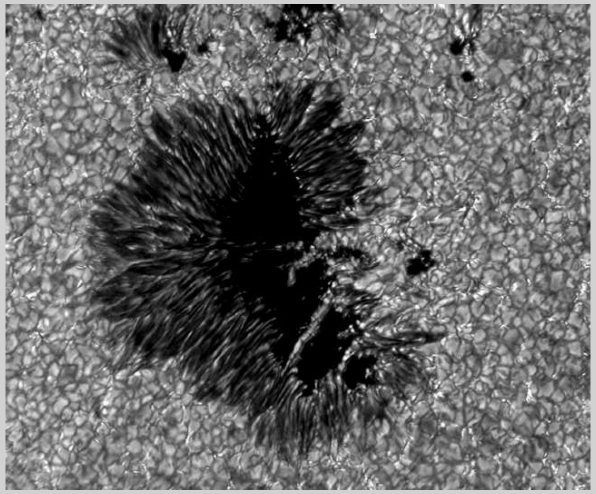
LONG-TERM VARIATION OF SOLAR MAGNETIC ACTIVITY

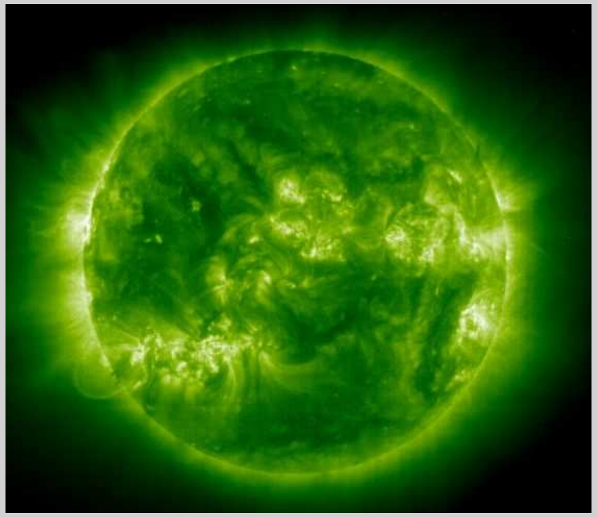
Antonio Ferriz Mas

Departamento de Física Aplicada
Facultad de Ciencias de Ourense
Universidad de Vigo
E-32004 Ourense (Spain)

Instituto de Astrofísica
de Andalucía (IAA)
Grupo de Física Solar
E-18080 Granada (Spain)







JOINT WORK OF:

José Ángel Abreu[1], Jürg Beer[1] & Antonio Ferriz-Mas[2]

[1]Swiss Federal Institute of Aquatic Science and Technology (EAWAG)
Dübendorf, Zürich
(Switzerland)

[2]Departamento de Física Aplicada
Universidad de Vigo
and
Instituto de Astrofísica de Andalucía
IAA/CSIC
(Spain)

COLLABORATION WITH:

Manfred Schüssler[3] and Dieter Schmitt[3]

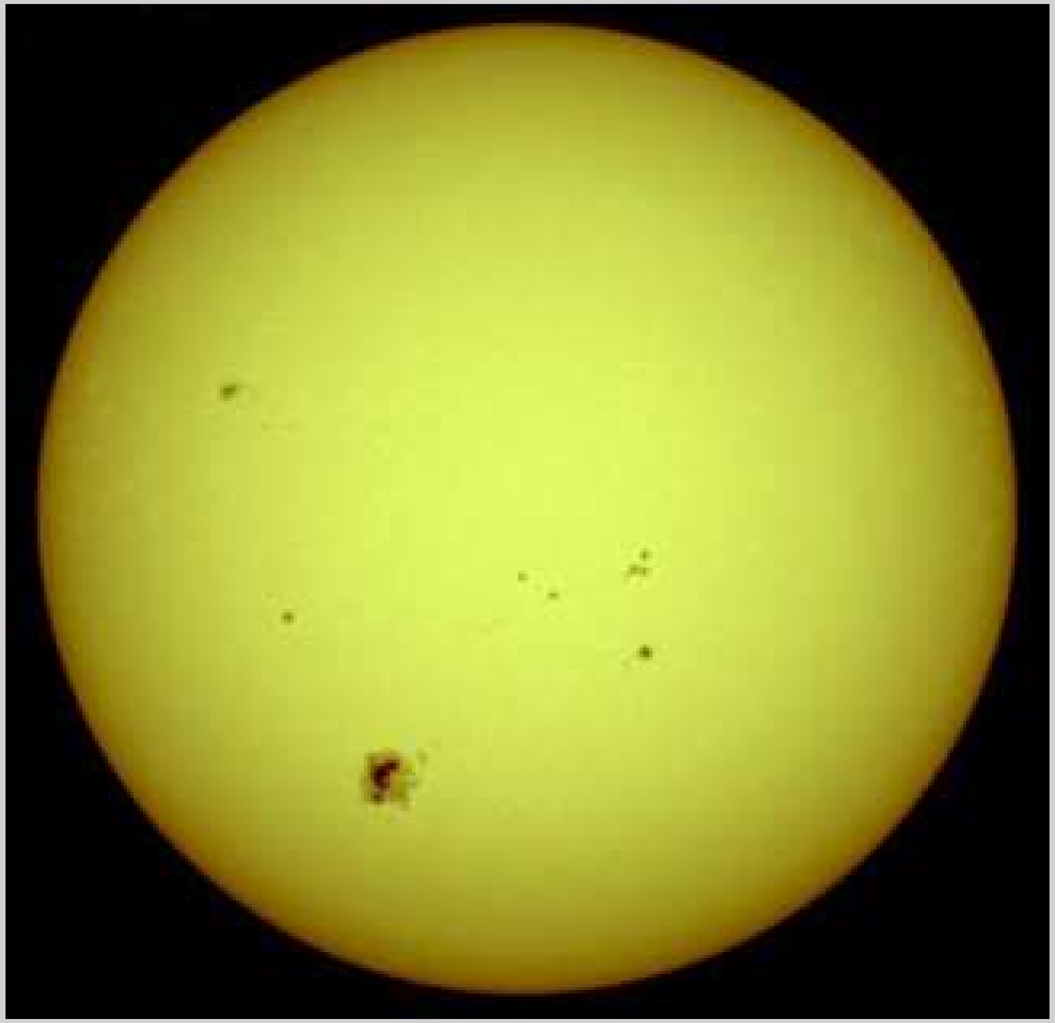
Max-Planck-Institut für Sonnensystemforschung
Lindau / Göttingen
(Germany)

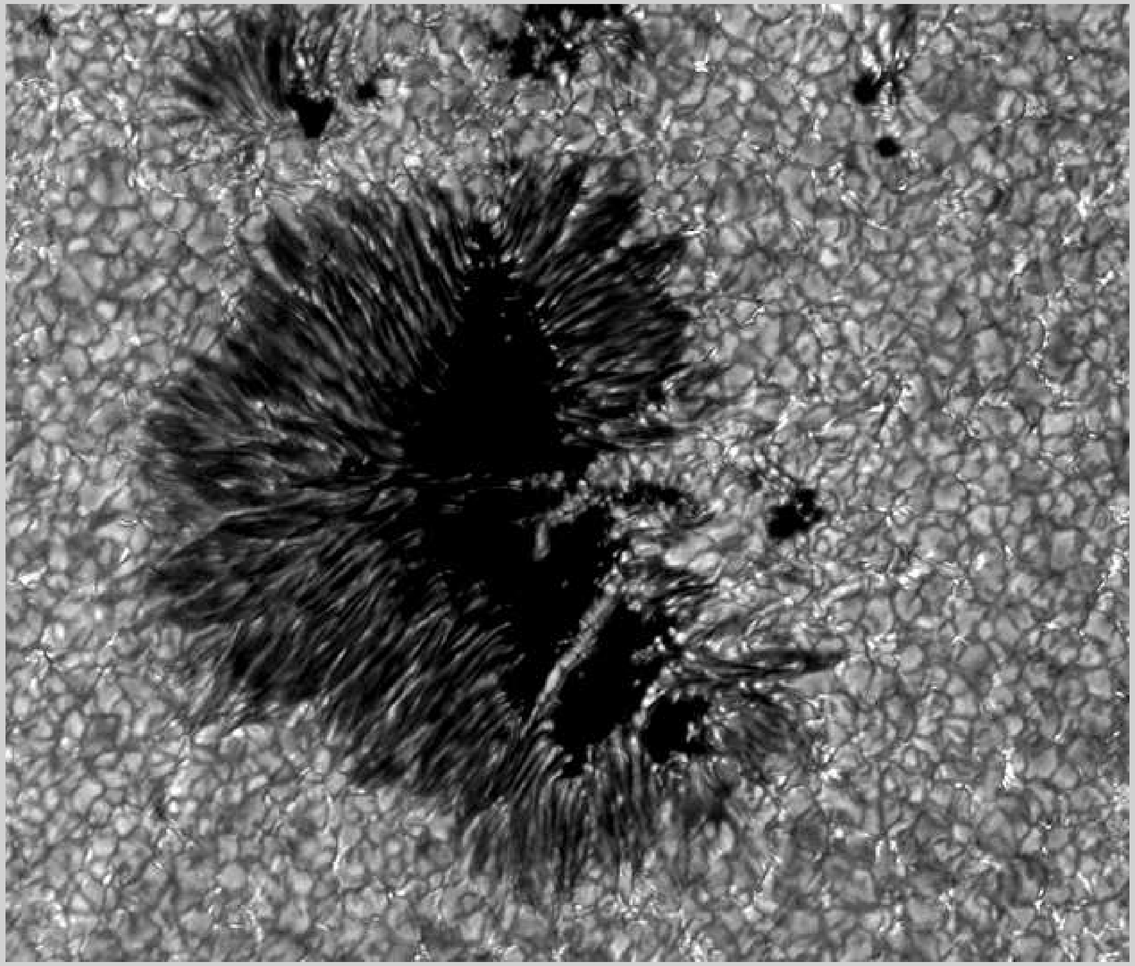
1. SUNSPOTS AND SOLAR MAGNETIC ACTIVITY

The magnetic field in the photospheric layers of the Sun does not occur in a homogeneous form, but in discrete concentrations of *intense* field.

(a) **Sunspots** are the most relevant manifestation of magnetic flux at the solar surface (albeit not the only only).

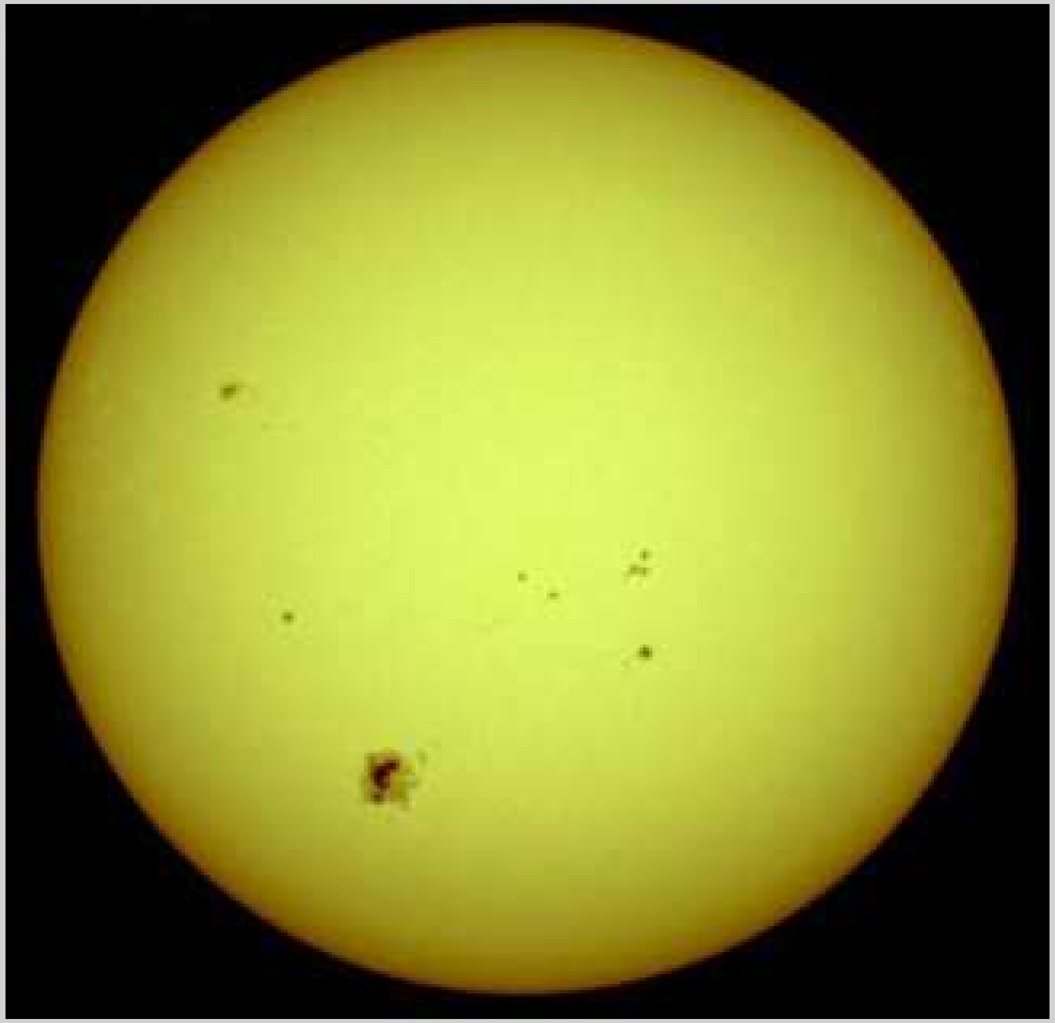
(b) **Sunspots** are cooler than the surrounding photosphere because they block part of the outgoing radiation. [Strong magnetic fields in a convective medium inhibit the energy transport].

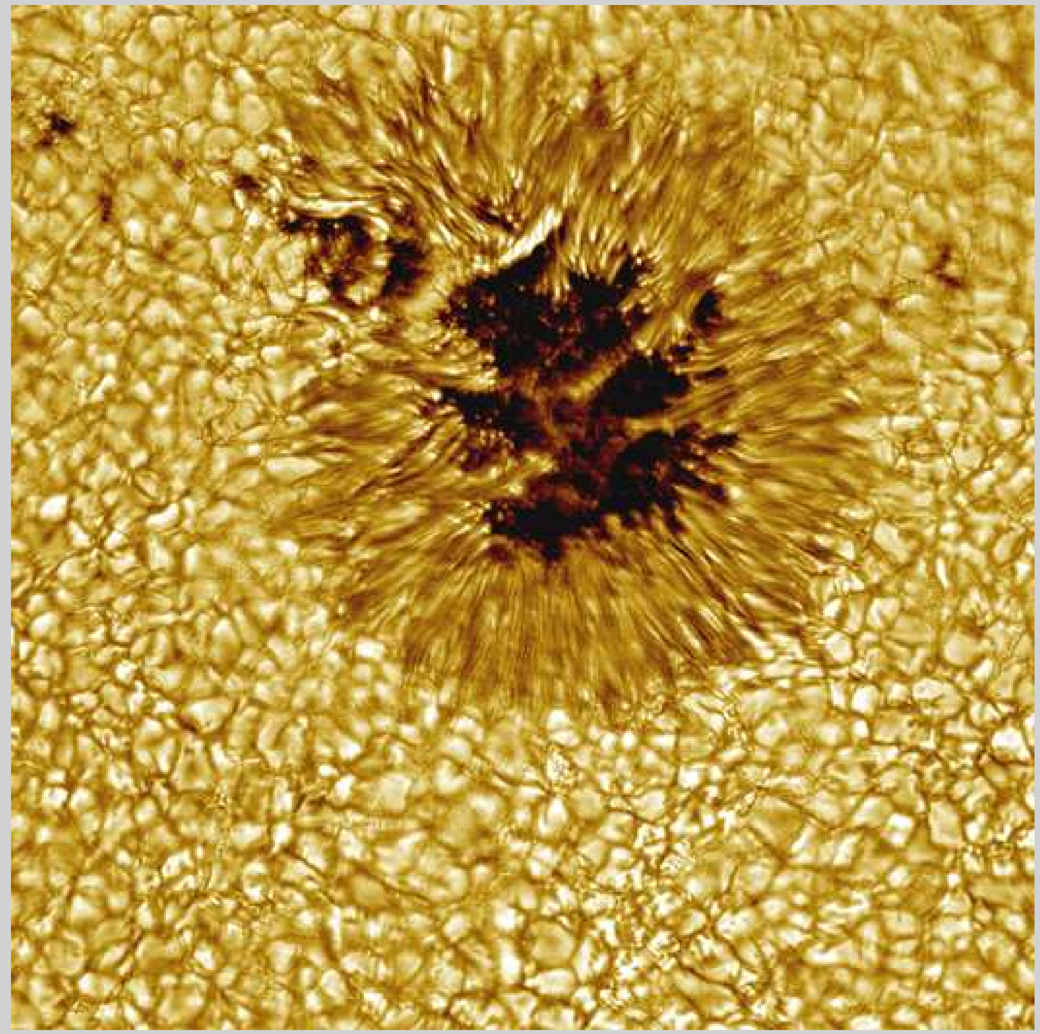




(c) One has traditionally associated "magnetic activity" with "number of sunspots", but recent data analysis of cosmogenic isotopes by Abreu et al. (2009) [see also Miyahara *et al.* (2007)] confirm that the solar magnetic activity did not disappear during the Maunder minimum and previous "grand minima"; i.e., it is wrong to identify "grand minimum" with "lack of dynamo action".

(d) In this meeting we are talking about sunspots because it seems now to be well established that the immediate cause of solar radiance variability is surface magnetism.





The photosphere

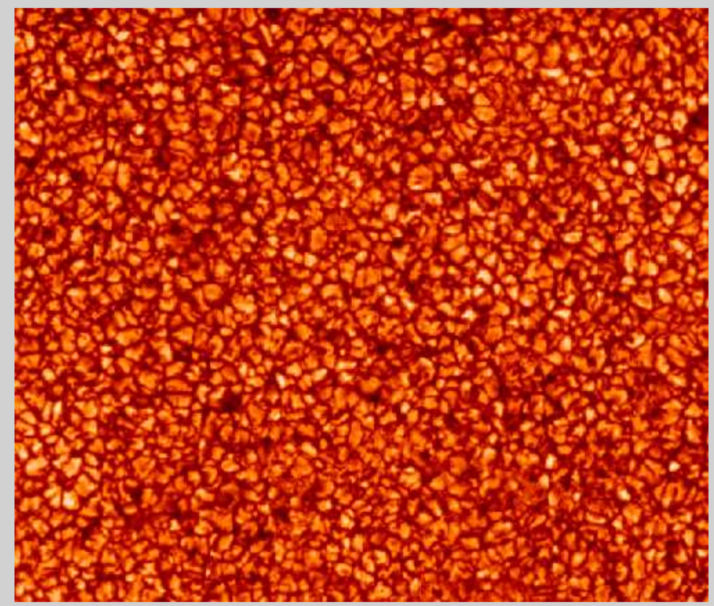
It is the visible surface of the Sun (matter is no longer *opaque* to the photons ⇒

- it is the thin layer from where most of solar light comes,
- it is where sunpots are seen.

It is a transition region of ca. **550 km** thickness between the convection zone and the tenuous *solar atmosphere.*

- Density $\simeq 10^{23}$ m^{-3} and temperature $\simeq 5\,800$ K .

Granulated structure of the photosphere:



Granules = convection cells (as seen from above).

Granules cover all solar surface (≃ 1 million granules at each time).

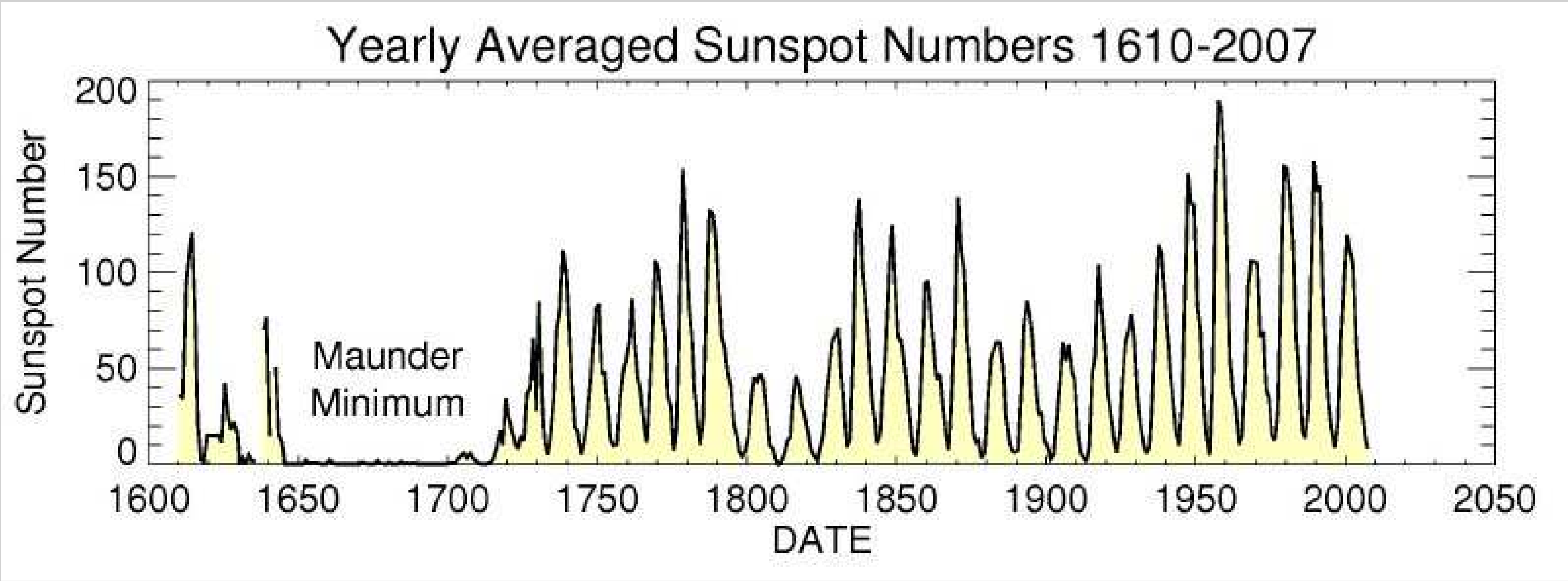
- Size: 700 – 1 500 km diameter.
- Upflow speeds: ≃ 0,4 km/s .
- Speed of horizontal spreading: ≃ 0,25 km/s .

[**Note:** Winds in hurricans may reach ≃ 0,1 km/s .]

- Average life time ≃ 8 minutes (exceptionally up to 15 min.)

SOLAR ACTIVITY CYCLE

The sunspot record since 1610 shows cycles of magnetic activity with an **irregular distribution of amplitudes** and with a period around 11 years. They were **interrupted by the Maunder minimum in the 17th century.**



REGULARITIES IN THE SOLAR ACTIVITY CYCLE

See e.g., *Solar MHD* by E. Priest

(a) Variation of the number of sunpots with a period of approx. 11 years ⇒ Activity cycle.

(b) Location of suspots on two activity belts above and below the equator.

(c) Equatorial drift of the activity belts in the course of the cycle. *(Butterfly diagram or Spörer's law)*. •

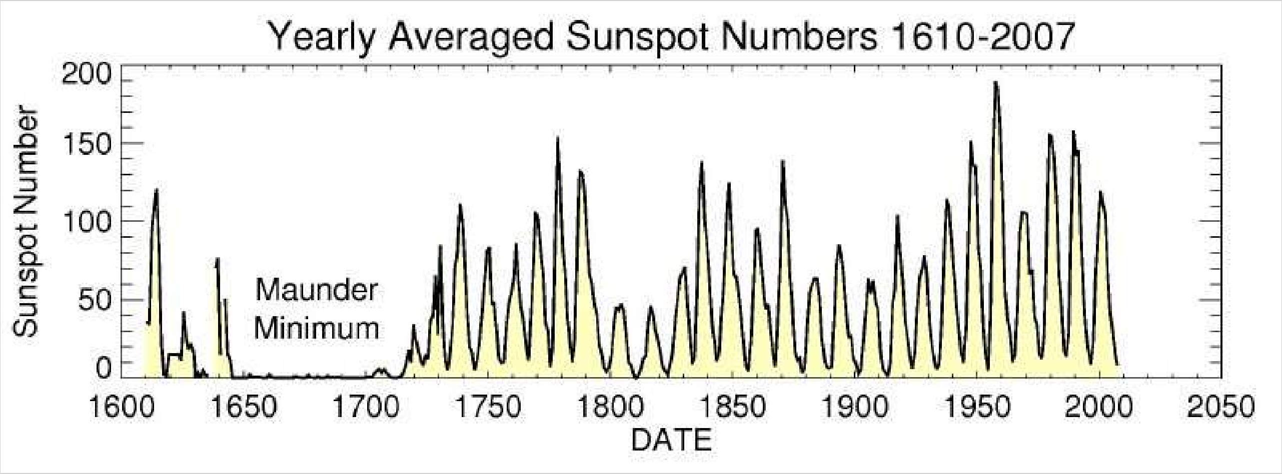
(d) Hale's *polarity rules*.

(e) Tilt angle of active region's main axis (by typically $10°$) with respect to the equator. •

(f) Systematische Asymmetrie zwischen den p- und f-Teilen bipolarer aktiver Gebiete.

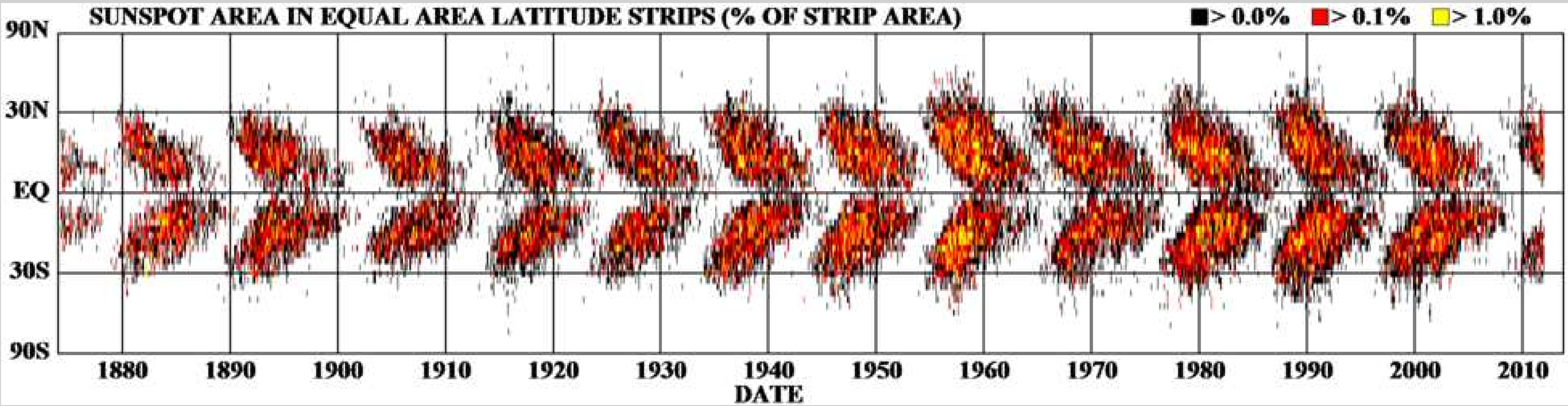
(g) Polarity reversals of the magnetic field near the activity minima ⇒ **cycle** of **22 years.**

(h) Long intervals ($\simeq 50-100$ years) with absence of activity in the form of sunspots (*grand minima*). Maunder minimum: period roughly from 1645 to 1715.



EQUATORIAL DRIFT over the cycle

Equatorial drift of the activity belts in the course of the cycle:
Butterfly diagram or Spörer's law.



2. COSMOGENIC ISOTOPES

Proxy indicators: Systematic observations of solar activity in the form of sunspots began with the invention of the telescope in 1609; before this date we must rely on proxy data as, for instance, the cosmogenic isotopes ^{10}Be and ^{14}C.

- ^{10}Be and ^{14}C are produced in the Earth's atmosphere by nuclear reactions involving particles from the *cosmic rays*. Both isotopes are produced in a similar way, but behave very differently after production.

- Cosmic rays are deflected by magnetic fields in the heliosphere and hence the production rates of these isotopes are modulated by changes in solar magnetic activity.

- The particles from the solar wind DO NOT produce cosmogenic isotopes (they are not energetic enough).

- Natural archives for ^{10}Be: ice in Greenland and in Antarctica.

- Natural archives for ^{14}C: tree rings.

- The *modulation potential* ϕ [first introduced by Gleeson & Axford (1968)] and is a measure of the role of the open solar magnetic field in deflecting cosmic rays.

- We have performed a composite of the solar modulation potential ϕ, reconstructed principally from the proxy record of cosmogenic ^{10}Be abundances in the GRIP ice core from Greenland. This composite of ϕ extends back for almost 10 000 years, showing many *grand maxima* and *grand minima* (Steinhilber *et al.* 2008).

Statistical analysis of the reconstructed composite of the solar modulation potential

- If we look at the record of sunspots, there is just on exceptional period of very low activity, the Maunder minimum.
- The ϕ composite used here extends back in time for almost 10,000 years, showing many grand minima and maxima, thus revealing that events like the well-known Maunder minimum are indeed quite normal in the history of solar activity.

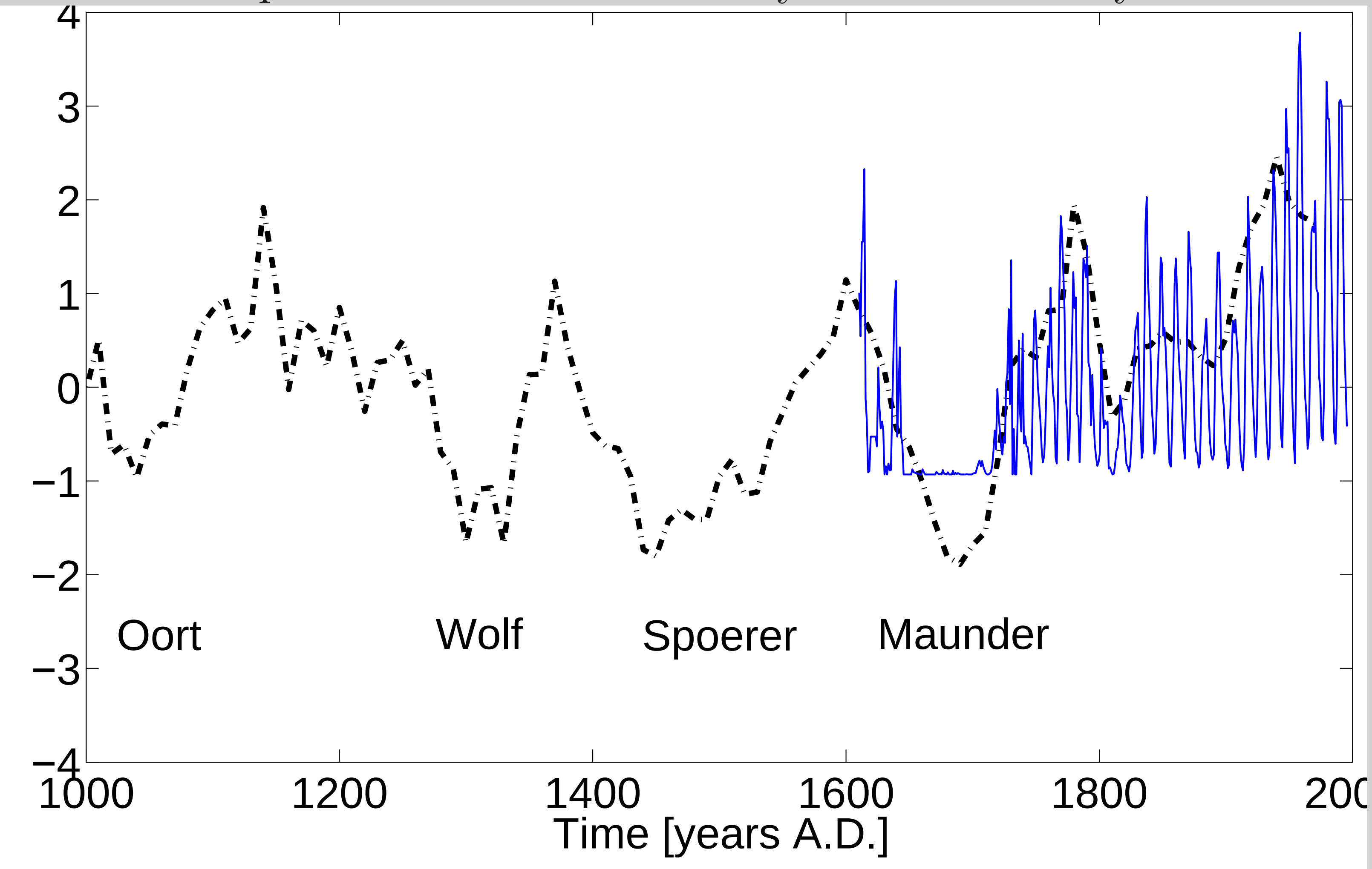


Fig. 1 shows a detail of the ϕ composite (dashed curve) together with the sunspot numbers R (blue solid line). We observe that the ϕ time series agrees quite well with the sunspot numbers: they display the same increasing trend, starting at the end of the Maunder minimum until present. The Dalton minimum around 1800, as well as the small minimum around 1900, are both well represented in our data. In this figure we see additional grand minima like Spörer's, Wolf's and Oort's, which lie outside the range of the sunspot data.

Features

- Looking at the whole set of ϕ data we see many grand maxima and grand minima (or Maunder minima).
- A frequency analysis of the solar modulation potential ϕ as well as of the ^{14}C–record reveals the presence of significant periodicities, namely around 2200 years (Hallstatt), 980 years and 200 years (de Vries).
- We see that "grand minima" tend to form clusters, for which no theoretical explanation is known.
- The information from the analysis of the compositer of ϕ place important constraints on dynamo models.

3. WHAT IS DYNAMO THEORY? WHAT IS THE DYNAMO PROBLEM?

The study of the growth and maintenance of magnetic fields in astrophysical bodies (planets, stars, compact objects, accretion disks, galaxies).

Astrophysical dynamos are governed by highly **non-linear** PDE's.

e.g., two of the governing equations are:

$$\rho \frac{D\mathbf{v}}{Dt} = -\frac{1}{\rho}\,\mathbf{grad}\, p + \rho\,\mathbf{g} + \frac{\mathbf{rot}\,\mathbf{B} \wedge \mathbf{B}}{4\pi} + \nu\,\nabla^2\mathbf{v}\,,$$

$$\frac{\partial \mathbf{B}}{\partial t} = \mathbf{rot}\,(\mathbf{v} \wedge \mathbf{B}) + \eta\,\nabla^2\,\mathbf{B}\,.$$

Even with simplifiying assumptions such as uniform density ($\Rightarrow$ no thermodynamics), uniform material coefficients η and ν, and uniform $\mathbf{g}$, the equations are non-linearly coupled through $\mathbf{v}$ and $\mathbf{B}$.

This renders impossible to find analytical solutions in all but the simplest cases and highlights the necessity of performing numerical simulations.

THE DYNAMO PROBLEM

Why is there a "dynamo problem"? Why not just solve numerically the set of MHD equations in a region (star, convection zone, accretion disk...) with appropriate boundary conditions?

(a) The Reynolds numbers (hydrodynamic and magnetic) are very large.

The large spatial dimensions of the system and the small value of the molecular viscosity cause strong nonlinear interactions in the flow, leading to a state of turbulent convection (magnetoconvection), for which no accepted theory exists nowadays.

- Massive numerical computations (e.g., Glatzmeier, Roberts, ...) reproduce successfully some important aspects of the geodynamo.

  But see papers by Ulrich Christensen in **Nature**, **454**, 1058–1059 and by Kageyama *et al.* in **Nature**, **454**, 1106–1109 (28 August 2008)... it cannot be claimed that the geodynamo problem has been solved.

- The situation is different for the Sun: so far, no numerical simulations reproducing –even qualitatively– the behaviour of the solar cycle have been performed. Why?

Main (but not only) difference with the geodynamo problem because stellar dynamos have much higher Reynolds numbers.

(b) The situation is further complicated by

- ◇ the effects of stratification (e.g., there are about 20 pressure scale heights in the solar convection zone),
- ◇ the existence of penetrative boundaries (at top and bottom),
- ◇ and the interaction of convection with rotation ($\Rightarrow$ differential rotation and meridional circulation).

Neither the spatial and temporal structure of the convection zone nor the profile of differential rotation can be directly derived from the basic equations of hydrodynamics.

(c) The hydromagnetic dynamo is hidden from direct observation.

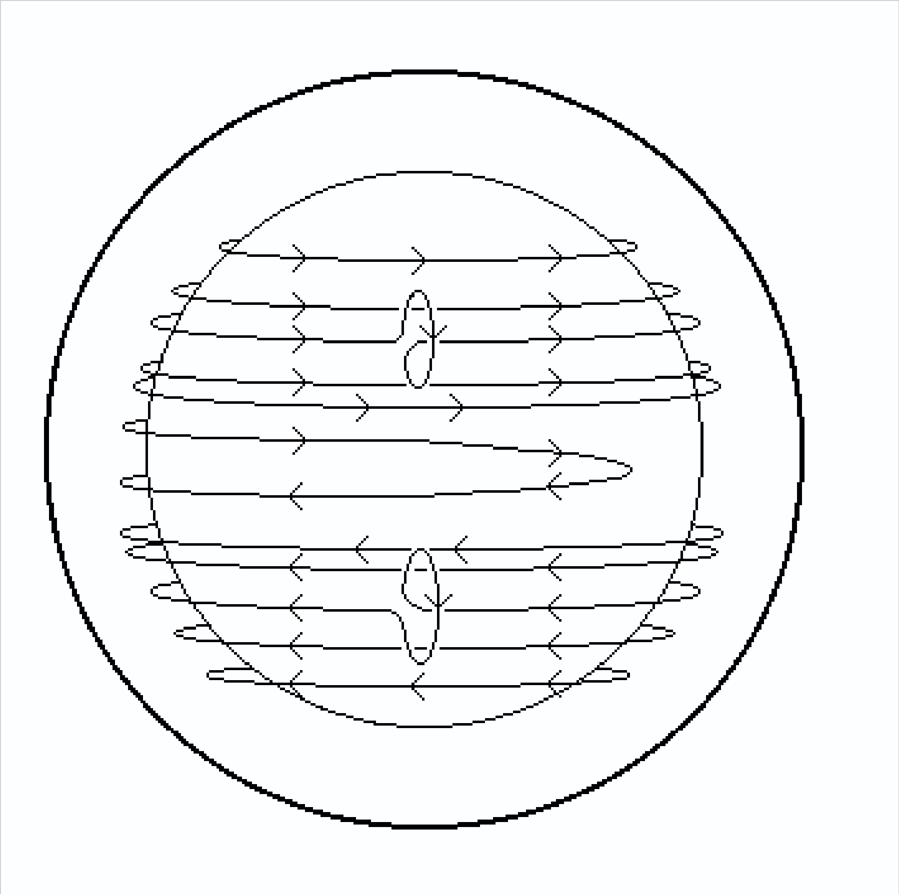
Only a shallow surface layer can be observed directly, while global oscillations (helioseismology) supply indirectly partial information from deeper layers.

The profile of differential rotation has been inferred from helioseismology and was in flagrant contradiction with the (at that time) generally accepted view that the solar interior rotates in cylinders.

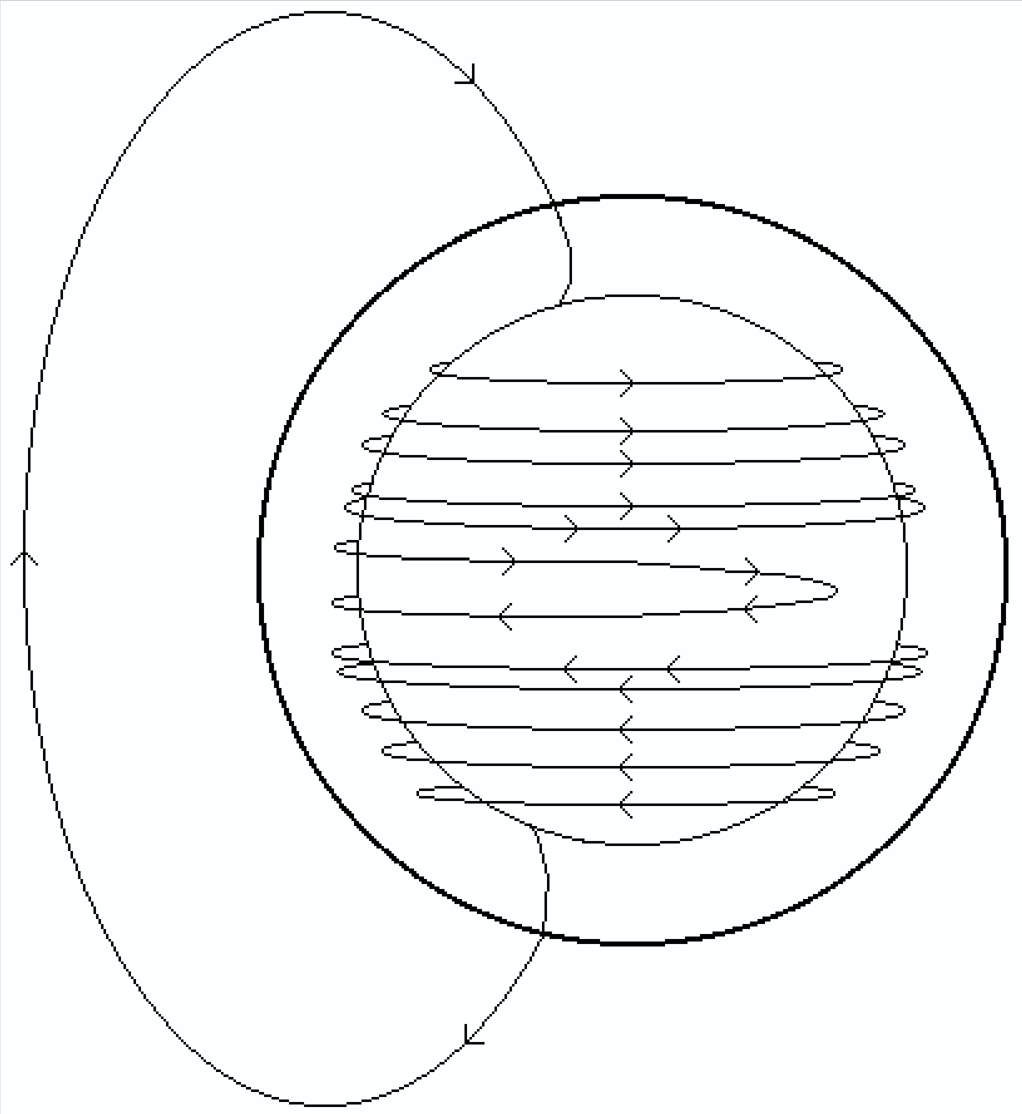
Conclusion: A complete theory encompassing in a self-consistent way the generation, structure, dynamics and evolution of magnetic fields in stellar convection zones does not exist yet.

THE CONVENTIONAL POINT OF VIEW

Most studies of stellar dynamos rely on **mean-field dynamo theory**, which is basically a kinematic approach [Parker, 1955].



***α*–effect:** Twisting of the magnetic field lines is caused by the **Coriolis force**. This is called the α–effect. The twist produced by the α–effect makes sunspot groups that obey Joy's law and also makes the magnetic field reverse from one sunspot cycle to the next (Hale's law).



Ω–effect: Magnetic fields within the Sun are stretched out and wound around the Sun by **differential rotation**. This is called the Ω–effect. The Sun's differential rotation with latitude can take a north-south oriented magnetic field line and wrap it once around the Sun in about 8 months.

The main ingredients of “mean-field dynamo theory” are
the α–effect, the Ω–effect and turbulent diffusitity.

- The Ω–effect is intuitive.
- The α–effect was introduced heuristically by Parker (1955).
- The enhanced diffusivity is a direct consequence of turbulence.

Later, Parker’s argument was given a formal basis by Steenbeck, Krause and Rädler within the framework of turbulent flows.

A first-principle demonstration of the $\alpha\,\Omega$–dynamo does not exist.

HOW TO APPROACH THE 'SOLAR DYNAMO PROBLEM'?

- Wait until computers become powerful enough to cope with it and in the mean time turn to other interesting (astro)physical problems.

- Perform numerical simulations in domains much smaller than the whole physical system (*e.g.*, magnetoconvection simulations *in a box*).

  Numerical simulations in magnetoconvection

  e.g., Nigel Weiss, Åke Nordlund, Robert ('Bob') Stein, Oskar Steiner, Manfred Schüssler, Axel Brandenburg, Fausto Cattaneo, Michael Proctor, Hiroaki Isobe, ...

- One way around the difficulties of tackling the full dynamo problem is to study a number of underlying processes in artificial isolation –usually in a simplified setting– in order to understand their fundamental physics and evaluate their relevance for the global problem.

  These fundamental processes are the building blocks of the full dynamo mechanism and must eventually be fitted together in order to understand it (e.g., storage of magnetic flux tubes, rise of unstable flux tubes through the convection zone, latitudes of emergence, tilt angles, flux emergence, etc.).

- Mean-field dynamo theory? This is the traditional approach (and very 'successful' in the sense that it is the source of countless papers).

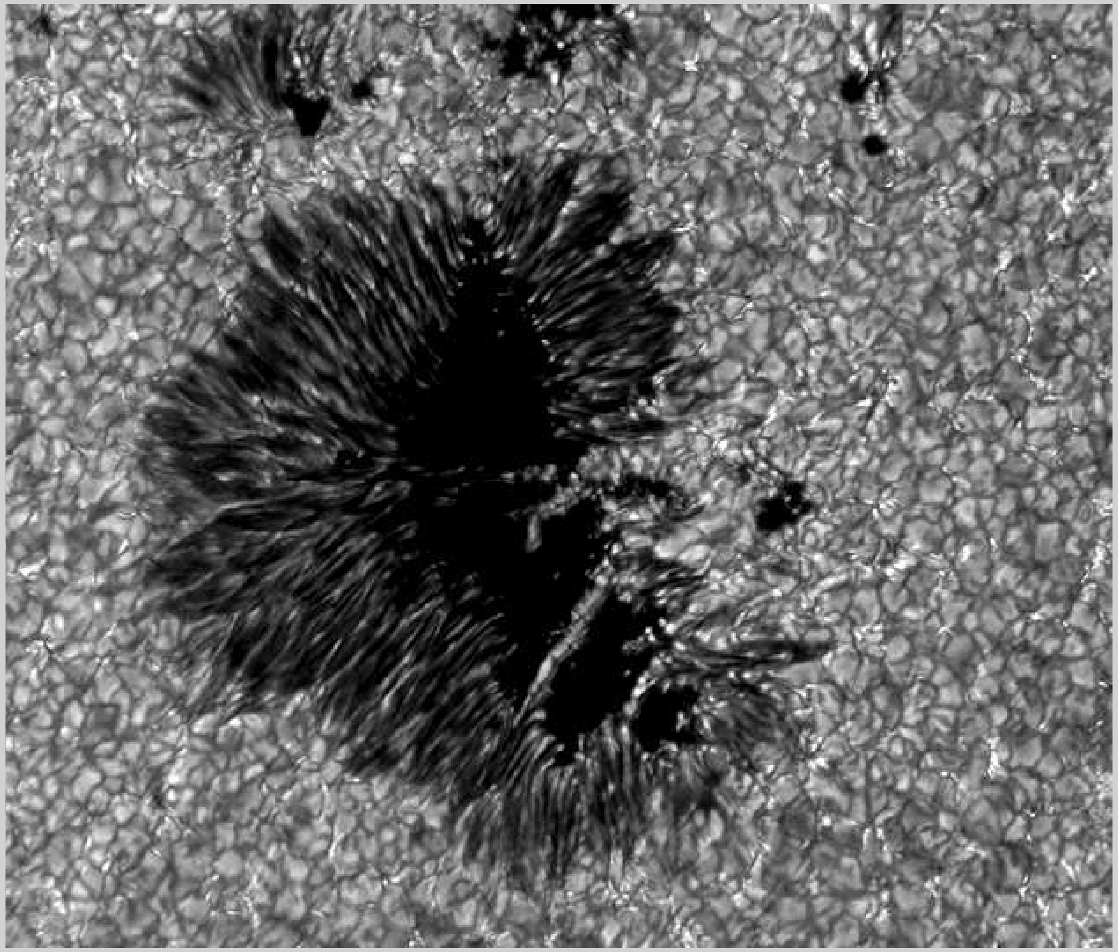
  Mean-field dynamo theory is called by some people "the realm of free parameters."

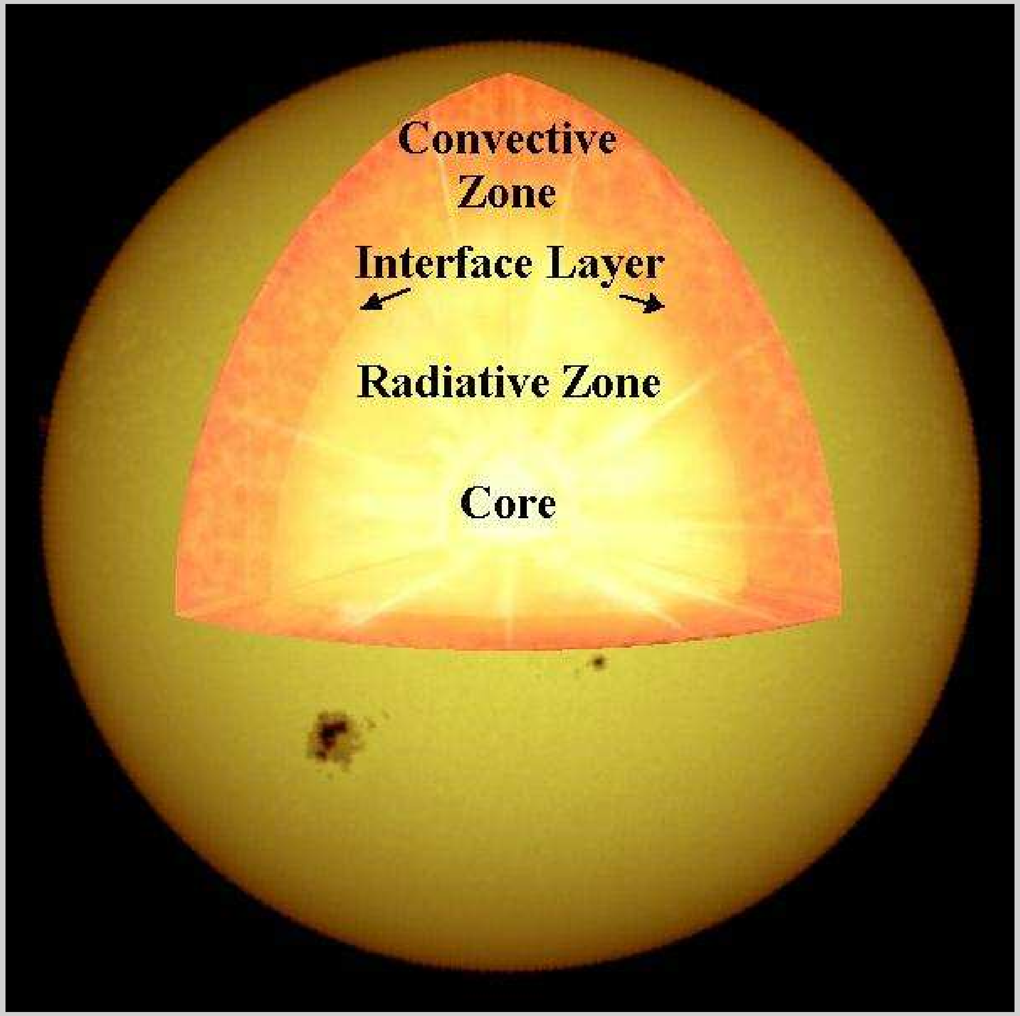
  In my opinion, "das Modell hat ausgedient".

4. LOCATION AND STRENGTH OF THE TOROIDAL FLUX SYSTEM

4.1. Where do sunspots come from?

- Bipolar active regions form by the emergence of large bundles of toroidal magnetic field from the convection zone into the photosphere.
- A number of **independent observational** and **theoretical** arguments support the hypothesis that the magnetic flux responsible for the cycle of sunspots is stored in a subadiabatic overshoot layer below the convection zone.





4.2 How strong is the field of the toroidal flux system?

Independent arguments indicate that the strength of the stored magnetic field prior to eruption is $\sim 10^5$ G [Moreno-Insertis, Schüssler & Ferriz-Mas 1992].

This value was first suggested by van Ballegooijen (1982), who considered hydrostatic flux tubes extending throughout the solar convection zone.

- Flux tubes with weaker fields would erupt at too high latitudes due to the action of the Coriolis force [Choudhuri 1989].

- Large fields are required
  - ◇ to account for the observed tilt angles of active regions [Caligari, Moreno-Insertis, Schüssler 1995].
  - ◇ to avoid excesive weakening of rising flux tubes (so-called *explosion of flux tubes*) [Moreno-Insertis, Caligari, Schüssler 1995].
  - ◇ to give account of the coherence of sunspots.

- Results of linear stability analysis of flux tube equilibria. [Ferriz-Mas & Schüssler, *Geophys. Astrophys. Fluid Dynam.* 1993, 1995].

Review of these arguments: **Convection zone flux tubes** (A. Ferriz Mas) in *Encyclopedia of Astronomy & Astrophysics, Nature Publishing Group, 2001.*

Reminder on the concept of SUPERADIABATICITY

- The superadiabaticity δ is simply a *dimensionless measure of the entropy gradient* in the vertical (or radial) direction when pressure is used as vertical (or radial) coordinate, *viz.*

$$\delta \stackrel{\text{def}}{=} \frac{p}{c_p} \frac{\mathrm{d}\,s}{\mathrm{d}\,p} \,.$$

where s is the specific entropy (*i.e.*, per unit mass) and c_p is the specific heat at constant pressure.

- A better known derivative of this kind is the *logarithmic temperature gradient*, ∇, defined as

$$\nabla \stackrel{\text{def}}{=} \frac{p}{T} \frac{\mathrm{d}\,T}{\mathrm{d}\,p} \,,$$

whereby temperature is considered as a function of pressure (i.e., the vertical coordinate z has been replaced by p).

In the case of a homoentropically stratified medium (inaccurately called 'an adiabatic stratification' in most of the astrophysical and geophysical literature), the corresponding dimensionless temperature gradient is denoted $\nabla_{\rm ad}$. It is straightforward to show that the superadiabaticity and the logarithmic temperature gradients ∇ and $\nabla_{\rm ad}$ are related by

$$\delta = \nabla - \nabla_{\rm ad} \,,$$

which is taken in most textbooks as definition of the superadiabaticity.

Application to a nonlocal mixing-length model OF THE SOLAR CONVECTION ZONE

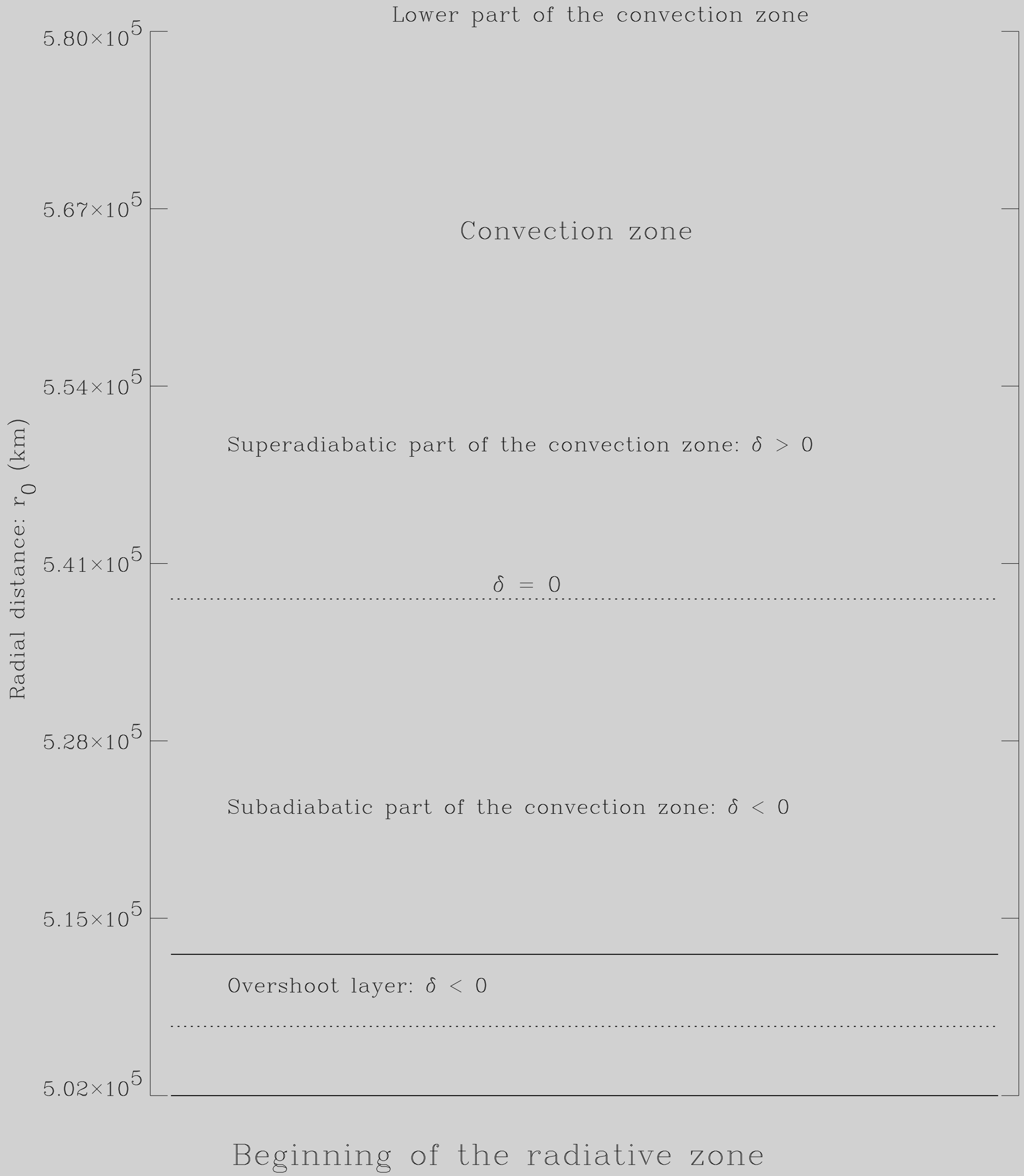
We use a solar model developed by M. Stix (cf. Skaley & Stix, 1991).

The model makes use of a non-local mixing-length treatment of the convection zone following the formalism of Shaviv & Saltpeter (1973), yielding a **consistently calculated overshoot layer** of about 10 000 km depth:

- Superadiabaticity becomes negative at $r = 5.3847 \cdot 10^5$ km.
- Bottom of the convection zone proper at $r = 5.1259 \cdot 10^5$ km (where $\delta = -4.2 \cdot 10^{-7}$).
- The radiative core begins at $r = 5.022 \cdot 10^5$ km (where $\delta = -1.4 \cdot 10^{-4}$).

IMPORTANT: The superadiabaticity $\delta = \nabla - \nabla_{\mathrm{ad}}$ becomes negative already in the lower part of the convection zone proper!

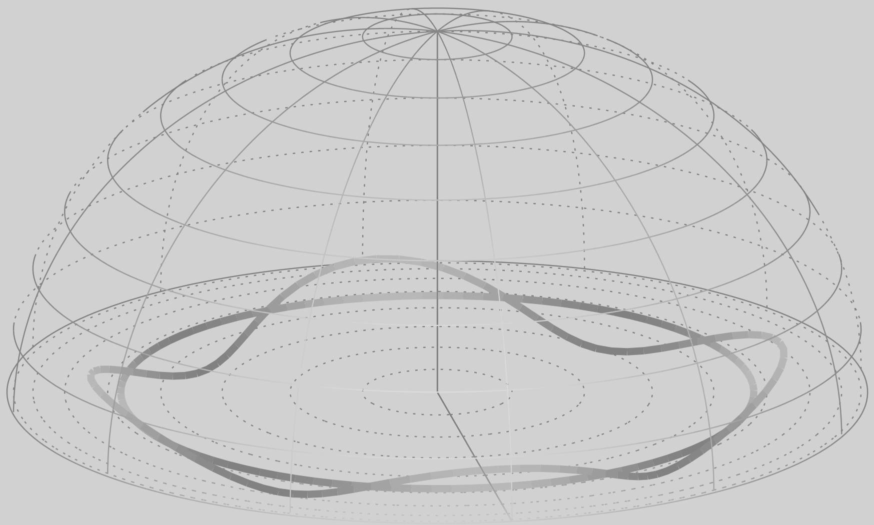
$\Rightarrow$ "subadiabatic layer" $\neq$ "overshoot region".



In the present model, the total extent of the sub-adiabatic layer is $\simeq 36\,000$ km, while the overshoot layer extends over $\simeq 10\,000$ km.

r_0 (km)	δ	p (dyn cm^{-2})
5.591e+5	2.001e-7	1.502e+13
5.385e+5	-1.781e-9	2.351e+13
5.126e+5	-4.232e-7	3.891e+13
5.073e+5	-9.071e-7	4.283e+13
5.023e+5	-1.170e-4	4.689e+13

4. BASIC IDEAS FOR A DYNAMO WITH MAGNETIC FLUX TUBES.



(a) Storage: Flux tubes stored in mechanical equilibrium in a layer of overshooting convection between the bottom of the convection zone and the radiative region.

The flux tubes are neutrally buoyant $\Rightarrow$ that mechanical equilibrium is a balance between **curvature force** and the **Coriolis force** due to a longitudinal mass flow in the tube.

For the Sun, it seems possible to store very intense (super-equipartition) flux tubes in mechanical equilibrium with $B_0 \gtrsim 0.9$–$1.1 \cdot 10^5$ G.

◇ Moreno-Insertis, Schüssler, Ferriz-Mas: *A&A* **264**, 686–700 (1992).

(b) Magnetic field amplification:

- differential rotation (Ω–effect) [*overshoot layer (possibly) = layer of strong shear*]
- and/or by other mechanisms (e.g.: explosion of 'weak' flux tubes in the C.Z. cf. Moreno-Insertis *et al.* 1995).

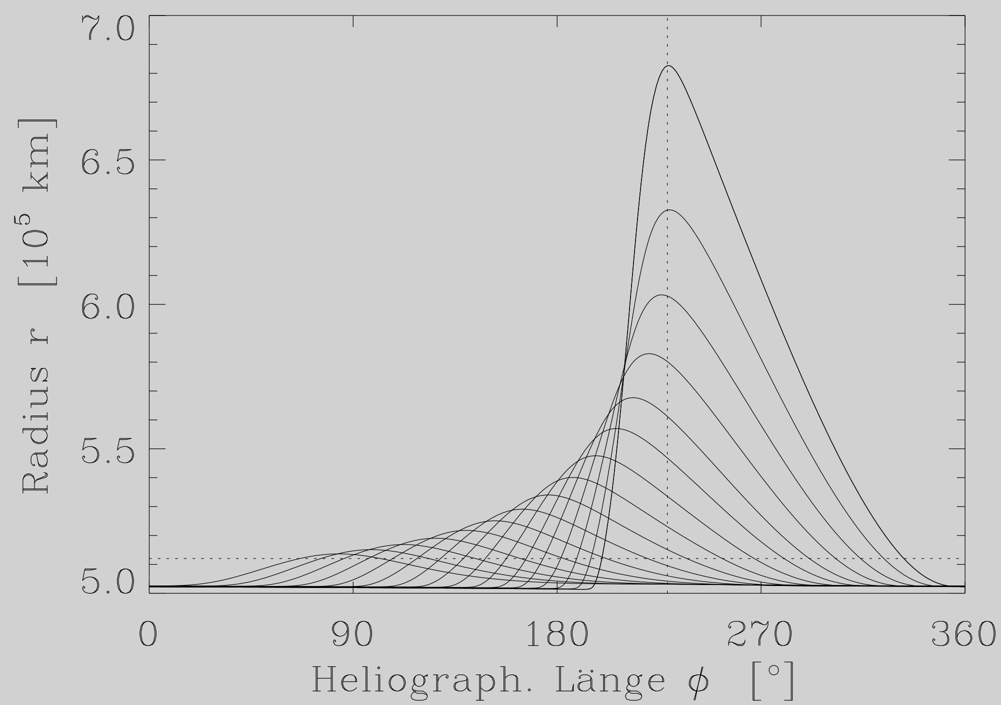
(c) Onset of instability:

For $B > B_c$ tube becomes unstable, enters the CZ and starts rising towards the surface.

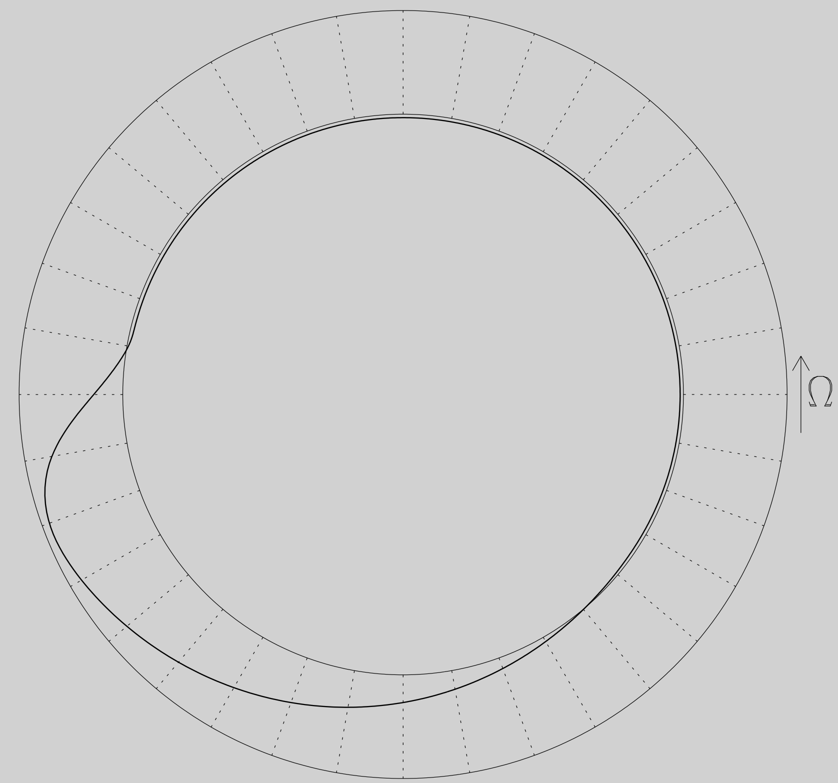
Critical value B_c depends on the stratification, latitude and angular velocity distribution.

(d) Numerical simulations of rise: (Peter Caligari, KIS, PhD Thesis).

- Eine Flußröhre mit $1.4 \cdot 10^5$ G im mechanischen Gleichgewicht wird gestört ⇒ Auftriebsinstabilität setzt ein und es bildet sich eine zweihöckerige *Seeschlange* (d.h., Eigenschwingung mit $m = 2$ ist am instabilsten).
- Eine der Schleifen steigt durch die Konvektionszone hinauf und bricht an der Oberfläche innerhalb von ca. 30 Tagen aus.
- Beim Ausbruch ⇒ Bildung einer aktiven Region.



Time evolution of an unstable magnetic flux tube from the bottom of the convection zone to the surface. The direction of rotation is towards increasing ϕ.



Projection onto the equatorial plane of the last stage of the rise of the flux tube.

GLOBAL PICTURE OF THE DYNAMO

Back in the 1990's Schüssler, Schmitt and Ferriz-Mas suggested in a series of papers a two-component model consisting of:

(I) A turbulent **weak-field dynamo** ($B < 10^4$) operating throughout the convection zone.

This turbulent dynamo can be probably described in its essentials by means of the classical mean-field approach (kinematic theory): α–effect due to combination of Coriolis force and cyclonic convection.

(II) A boundary layer, **strong-field dynamo** ($B \gtrsim 10^5$) located in the overshoot layer with strong (super-equipartition) fields concentrated in isolated flux tubes. Depth $\simeq 10^4$ km.

Each dynamo is responsible for different aspects of the solar activity cycle:

(I) The weak (diffuse) dynamo generates a more irregular field.

Such fields would not disappear during a grand minimum and probably maintain a reduced level of activity.

(II) The strong (flux tube) dynamo is ultimately responsible for the solar activity cycle of sunspots.

Grand minima would appear when this dynamo ceases to work.

6. CONCLUSIONS, OPEN QUESTIONS AND PERSPECTIVES

- A stellar/planetary dynamo is a complex process by which kinetic energy, gravitational potential energy and/or thermal energy are converted into magnetic energy.

- The full dynamo problem is a difficult one... but this is not a surprise.

  The full dynamo process takes place in a region comprising approx. 20 pressure-scale heights.

  Why should the solar dynamo problem be easier to deal with than the weather problem or the climate problem?

  Parallelism between the convection zone and the atmosphere (troposphere): turbulent convection, differential rotation, meridional circulation, waves, penetrative boundaries...

- One way around the difficulties of tackling the global dynamo problem is to study a number of underlying processes in artificial isolation –usually in a simplified setting– in order to understand their fundamental physics and evaluate their relevance for the global problem.

  These fundamental processes are the building blocks of the full dynamo mechanism and must eventually be fitted together in order to understand it (e.g., storage of toroidal magnetic flux, rise of unstable flux tubes through the convection zone, explosion and possible intensification of flux tubes, latitudes of emergence, tilt angles, flux emergence, etc.).

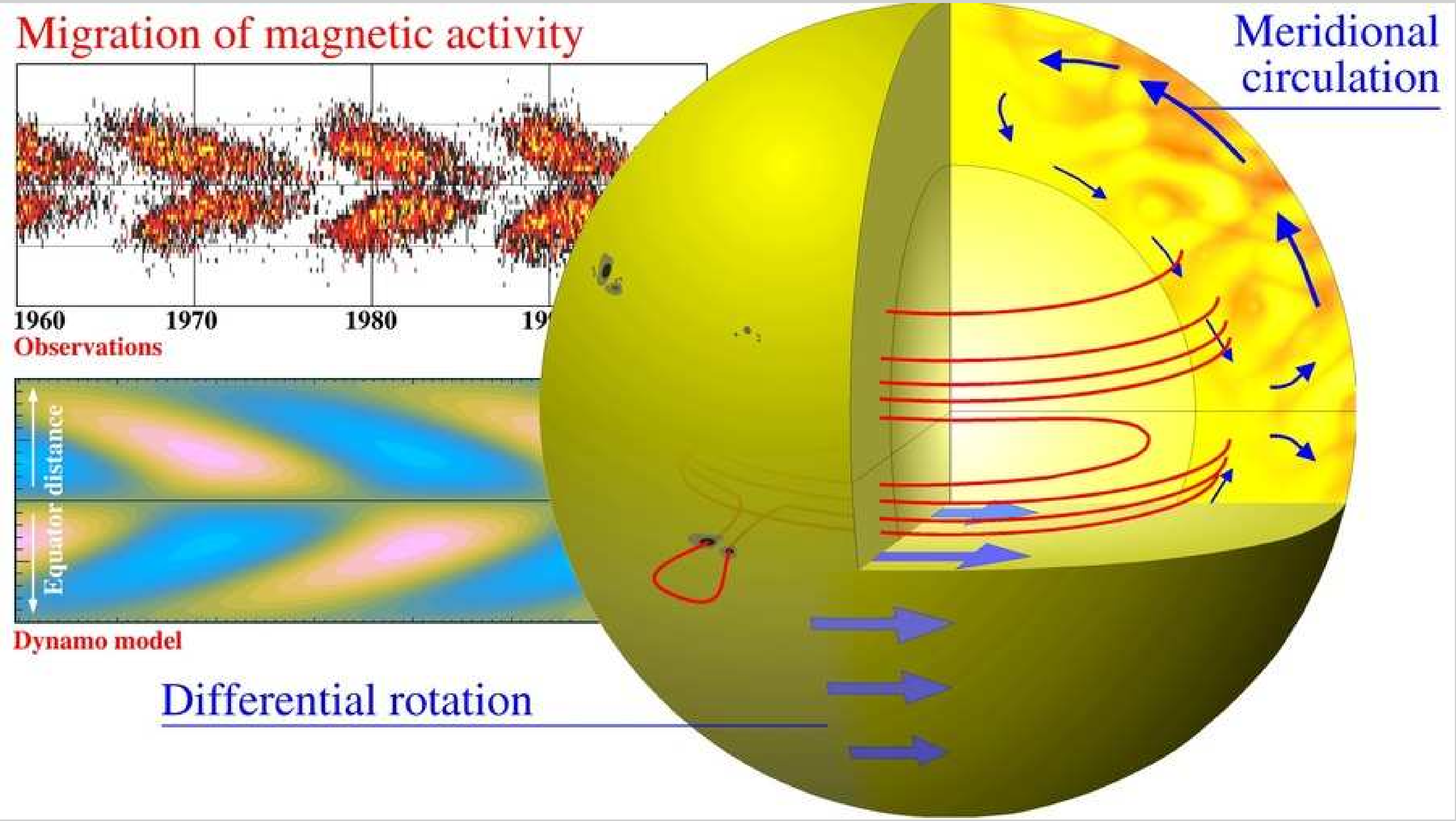
- The mean-field dynamo approach (often called "mean-field dynamo theory" or also "mean field electrodynamics") has been the traditional approach to the dynamo problem.

  It has been very successful... in the sense that it is the source of countless papers.

- New ideas/approaches are needed after more than 50 years of playing with free parameters.
- The solar magnetic activity shows a number of periodicities (starting with the 11-year cycle) that have not been explained.

  A possibility is that all or some of these periodicities are the result of complex flow patterns in the solar interior.

  Meridional circulation -long forgotten- and possibly the **excitation of the overshoot layer by gravity waves** probably play a key role in determining these periodicities.



- The key factors determining the storage capacity of the overshoot layer is the superadibaticity, $\delta = \nabla - \nabla_{\mathrm{ad}}$, which is a dimensionless measure of the stratification of the specific entropy in a medium.
- Any effect that can change δ can influence the magnetic storage capacity of the tachocline.
- The maximum field strength of a flux tube that can be stored at a given latitude in the overshoot layer depends very sensitively on the value of δ, which is of order 10^{-4} to 10^{-5}. Tiny variations in δ may decide whether a flux tube becomes unstable at $2 \cdot 10^4$ G or at 10^5 G. This makes a great difference, because flux tubes that do not reach a strength close to 10^5 G before entering the convection zone cannot reach the solar surface as a coherent structure and form sunspots.