UCL

Solar Activity and the magnetic cycle

Lidia van Driel-Gesztelyi

(1) University College London, Mullard Space Science Laboratory, UK
(2) Observatoire de Paris, LESIA, France
(3) Konkoly Observatory, Budapest, Hungary

E-mail: Lidia.vanDriel@obspm.fr; lvdg@mssl.ucl.ac.uk

UCL

Outline

- **Global properties of the solar cycle**
  - sunspots and groups - magnetic flux
  - irradiance
  - T_e, R
  - longer timescales
  - cycle 23 and 24 - actualities
  - methods of cycle forecast

- **Spatially resolved properties**
  - main laws: - Hale-Nicholson's law
    - Spörer's law (butterfly diagram)
    - Joy's law
  - torsional oscillations
  - meridional flow
  - hierarchy in flux emergence: sunspots, active regions, nests
  - characteristics of flux emergence: inclination, tilt, twist
  - evolution of active regions
  - disappearance of flux
  - global flux budget of the Sun

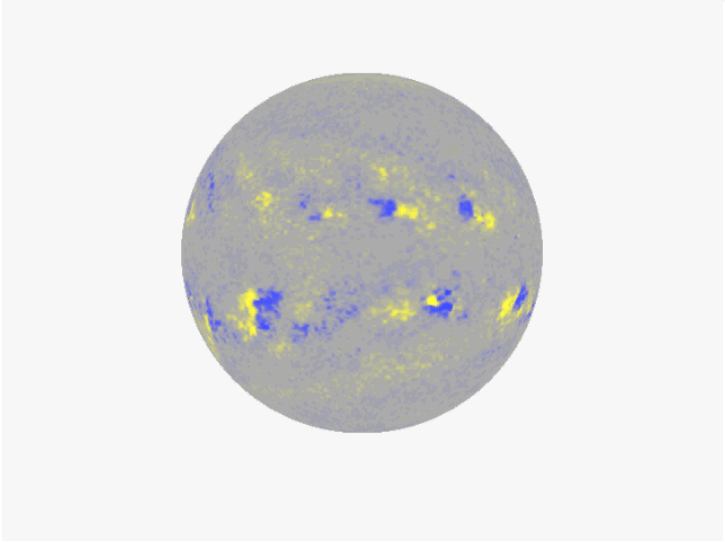
UCL

Main sources

- C.J. Schrijver & C. Zwaan: Solar and stellar magnetic activity Cambridge Astrophysics Series 34, Cambridge Univ. Press, 2000.
- K.L. Harvey: 1993, PhD Thesis, Univ. Utrecht, The Netherlands
- Review talks at the ITP Santa Barbara Workshop on Solar Magnetism in Jan. 2002 by D. Hathaway, D. Rabin, G. Fisher
- Review talks at the ISSI workshop on "Solar Magnetism" in Jan. 2008 and at the meeting "Cycle 24" in Dec. 2008 in Napa, California by D. Hathaway.

UCL

Movie of two full magnetic cycles



Movie by
D. Hathaway

Blue: negative; yellow: positive magnetic polarity. Kitt Peak Observatory magnetograms

UCL

Solar cycle - history

- 1843: The 11-year cyclic variation in the number of sunspots was discovered by H. Schwabe.

- 1849: R. Wolf started to tabulate daily sunspot numbers
  + started to reconstruct sunspot behaviour back to the beginning of the 17th century.

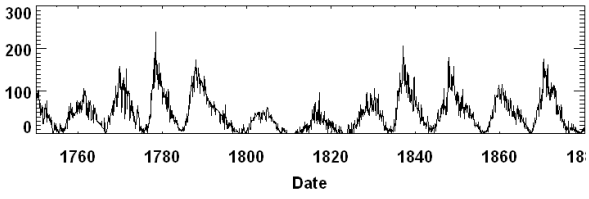
Wolf's sunspot index or Zürich relative sunspot number is defined as:

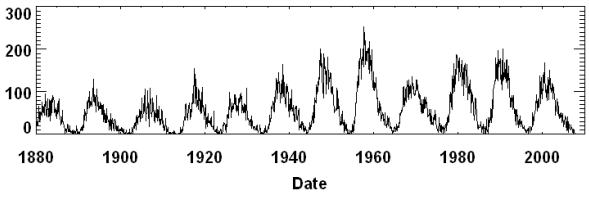
$$R = k(10g + f)$$

where **f** is the **number of individual spots** visible on the disk, **g** is the **number of sunspot groups** (ARs with spots) and **k is a correction factor** to adjust for observers, telescopes and site conditions.

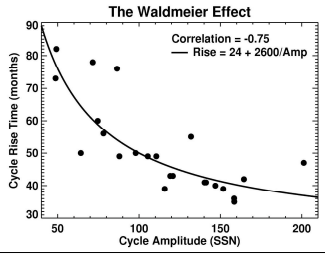
UCL

The '11-year' cycle



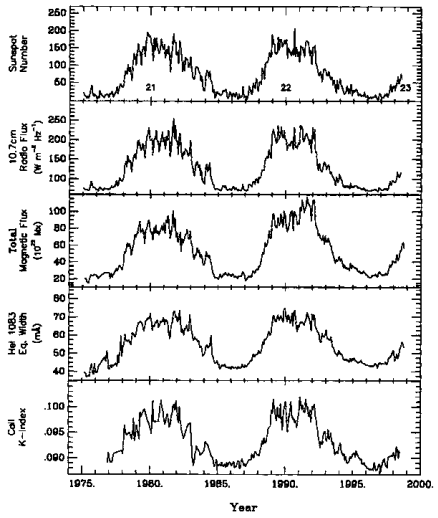


- amplitude variations with factor of 3
- length: 8-15 yr
- mean length: 11.1 yr
- asymmetry in rise and decline - strongest for high-amplitude cycles (Waldmeier effect)



UCL

Activity indices for cycles 21 & 22 averaged over a Carrington rotation



← Zürich relative sunspot number
R ≡ k (10g + f)

← 10.7 cm radio flux (DRAO)
(see next slide)

← Total magnetic flux (10^{22} Mx)
(Kitt Peak; most physical)

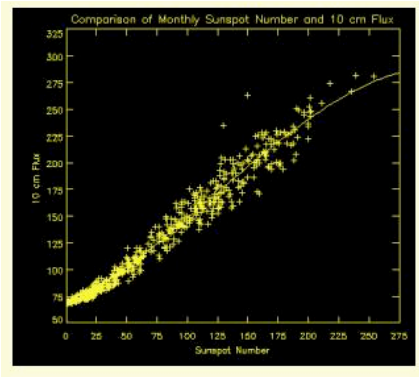
← He 10830 Å equivalent width

← Ca II K index

(plot by K. Harvey)

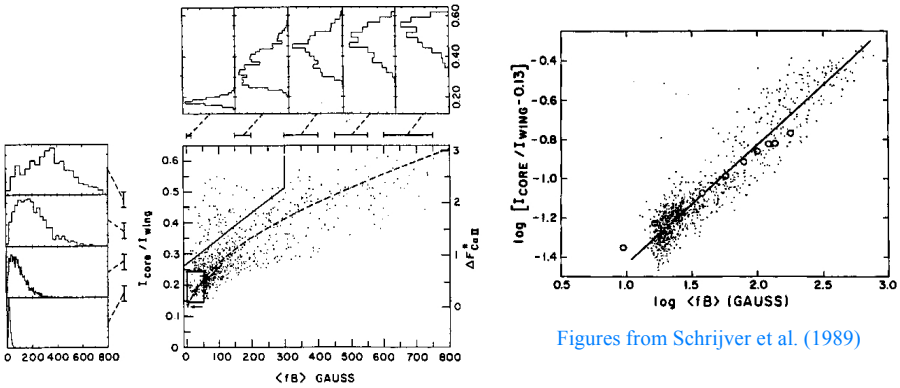
UCL

The 10.7 cm radio flux-sunspot number dependence



Dominated by diffuse thermal free-free radiation above ARs and gyroresonance emission from electrons in magnetic fields above sunspots.

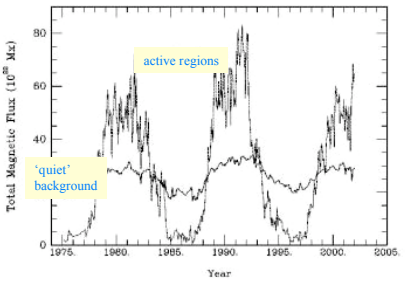
The Call K line-core intensity ratio I_{core}/I_{wing}



Figures from Schrijver et al. (1989)

Emission from plages and network - the ratio of the core intensity to that of the more stable background gives a robust indication of the chromospheric activity. The ratio's close correlation with magnetic field strength makes it act as a "magnetometer", sensitively following the cycle.

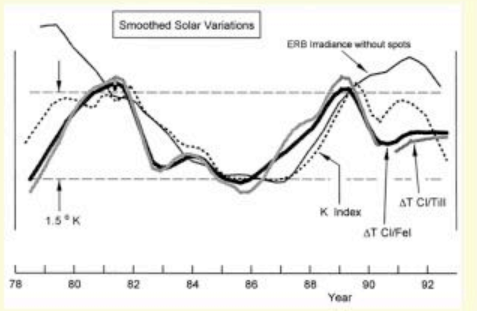
Cycle dependence of strong and weak magnetic flux



Harvey 1993

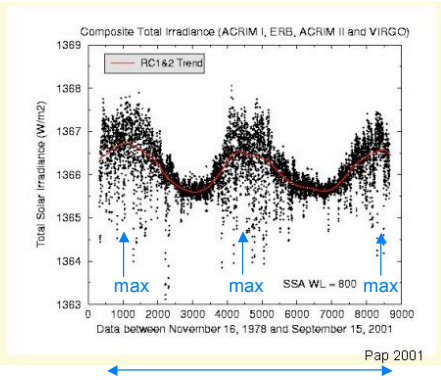
The cycle amplitude is much smaller for the weak 'background' field.

Effective temperature

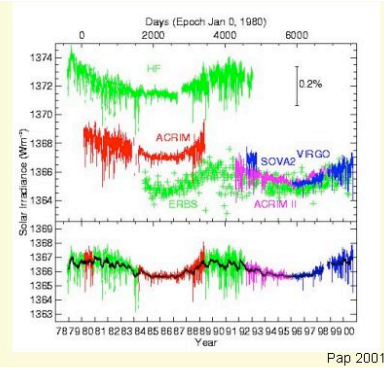


The C I 5380 to Fe I 5379 and to Ti II 5381 lines in the solar flux spectrum are robust indicators of the effective temperature. T_{eff} appears to vary with other cycle indicators with an amplitude of the variation of 1.5 K ± 0.2 K, similar to but slightly less than the range of the sunspot-corrected irradiance. (Gray and Livingston, 1997)

Total solar irradiance



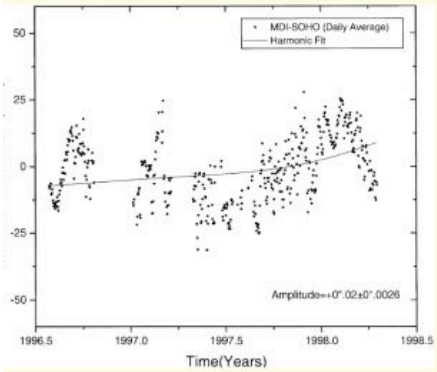
23 years



Irradiance changes are due to the competition between dark sunspots and bright faculae (Wilson & Hudson, 1991), in which the faculae win: they over-compensate the effect of sunspots. The total solar irradiance changes by 0.2% from cycle maximum to minimum (Frölich, 1994).

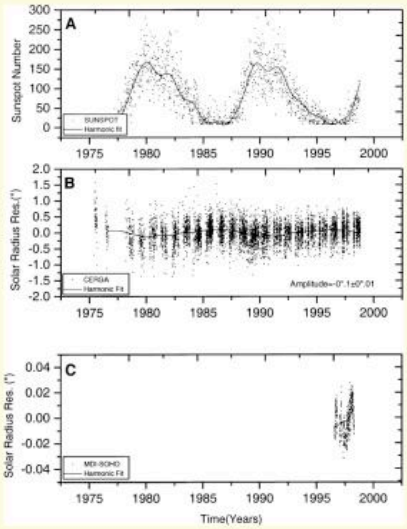
UCL

Solar radius



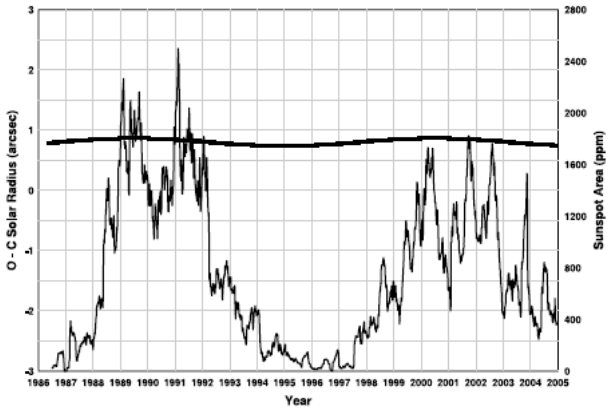
The residual linear variation in the Sun's radius is 8.1 ± 0.9 mas yr^{-1}.

Emilio et al. 2000



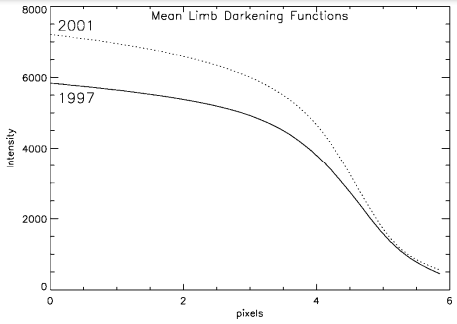
UCL

Solar radius



Chapman et al. 2008:
Using ground-based full-disk 672.3 nm images obtained a peak-to-peak amplitude of 0.136''±0.01'' in the N-S geocentric solar radius; larger at solar maximum during cycles 22 & 23.

Solar radius

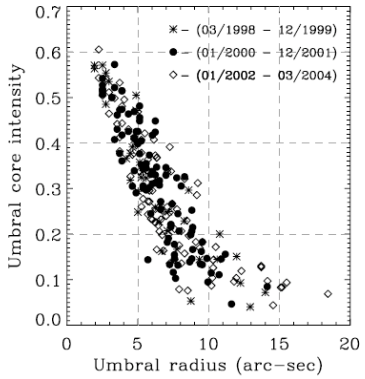


Emilio et al. 2007:

The apparent acoustic equatorial radius decreases at solar maximum, while the optical radius increases ⇒ the outermost solar atmosphere must expand.
The outer solar atmosphere expands non-homologously during the cycle, possibly due to a changing turbulent pressure. This leads to changes in the limb darkening function.

Sunspot umbral brightness

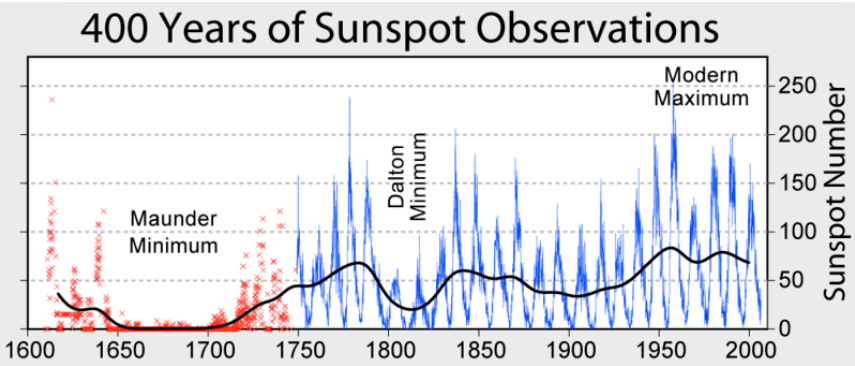
Albregtsen & Maltby (1978) found a slight but significant (0.15) cycle variation in the umbral brightness of spots:
They showed that the first sunspots of the cycle are the darkest and the umbral brightness increases
linearly with time. They showed that this is not a latitude effect, but a true temporal evolution. However, their sample contained only 13 spots…



This result was not confirmed by Mathew et al. (2007), who found brightness variation only with the umbral radius but no cycle dependence…

UCL

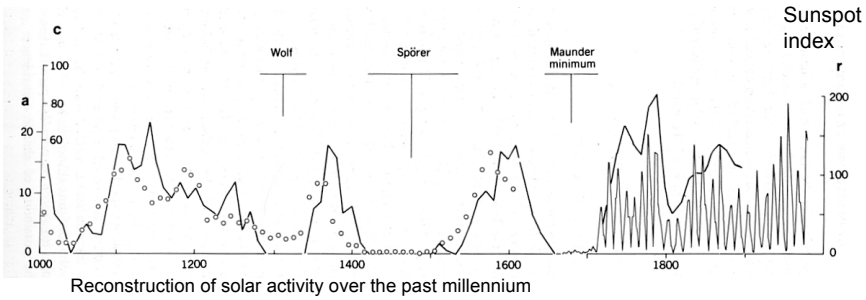
Longer-term trends



Early records of sunspots indicate that the Sun went through a period of inactivity in the late 17th century (Maunder, 1890). Very few sunspots were seen on the Sun from about 1645 to 1715. Although the observations were not as extensive as in later years, the Sun was in fact well observed during this time and this lack of sunspots is well documented (Hoyt and Schatten). There is evidence that the Sun has had similar periods of inactivity in the more distant past. More recent low-activity period was the Dalton minimum (1790-1830).

UCL

Recent Grand Minima



Reconstruction of solar activity over the past millennium

- from telescopic sunspot observations (light solid line, starts ~ 1650)
- proxy of sunspot number from C^{14} data (heavy solid line left, index c)
- northern hemisphere aurorae (circles, in sightings per decade, index a)

Charateristics:

- Gleissberg cycle (presence of a superposed 80-100 years periodicity)
- Grand minima (Wolf, Spörer, Maunder)
- High activity period in the 12th-early 13th century

→ **Effects on the terrestrial climate**

UCL

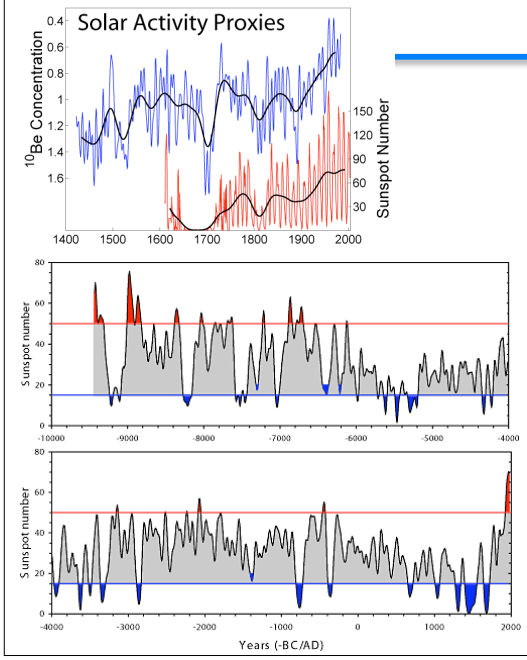
Maunder minimum - little ice age

Frozen Thames, 1677 (unknown author, London)

Maunder minimum (1645 - 1715): This period of solar inactivity also corresponds to a climatic period called the "Little Ice Age" when rivers that are normally ice-free froze and snow fields remained year-round at lower altitudes.

UCL

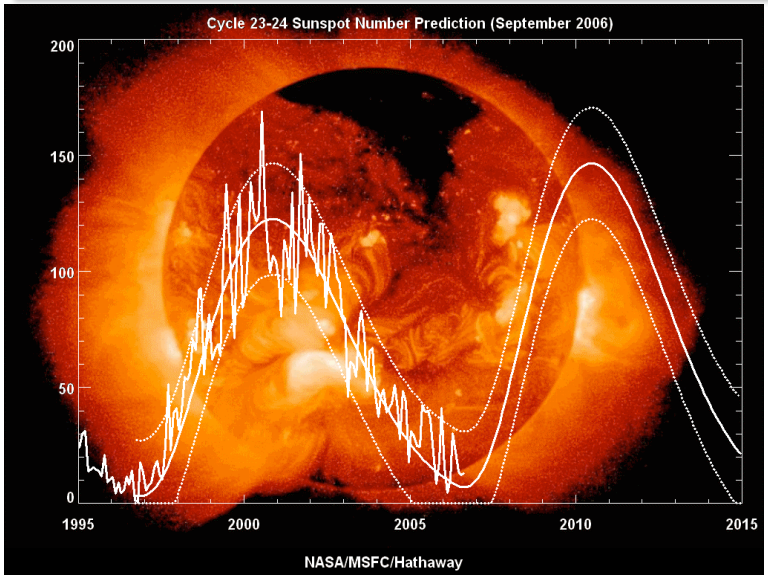
Long-term cycle data



^{10}B in ice cores and ^{14}C in tree rings are produced in the Earth's atmosphere by galactic cosmic ray collisions which are anti-correlated with the sunspot cycle. 10000-year reconstructions of solar activity from these records indicate about 15 Grand Maxima and 25 Grand Minima over this time period (one in every 400 years). Note that presently we see the end of a Grand Maximum.
(Hathaway, 2009)

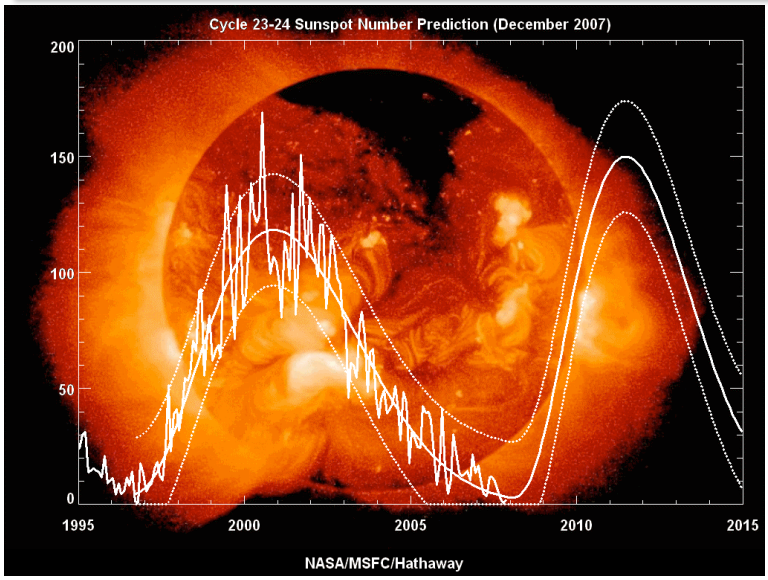
UCL

Present and next cycle



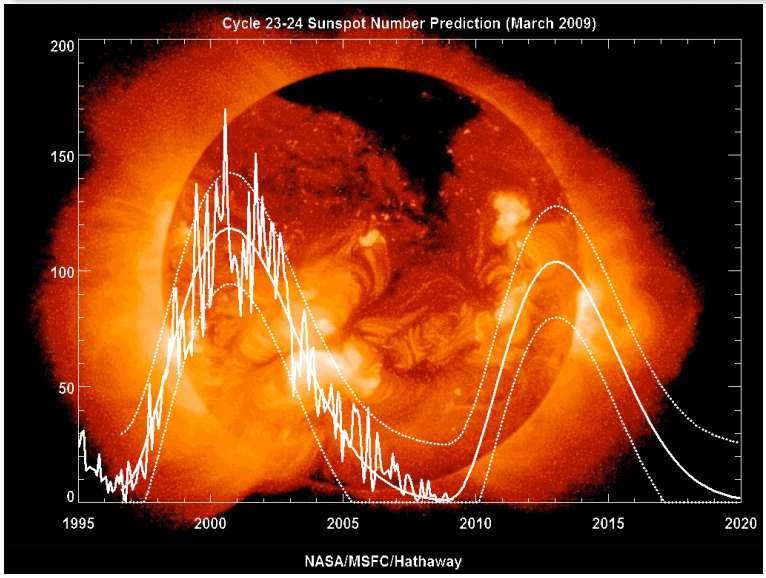
UCL

Present and next cycle



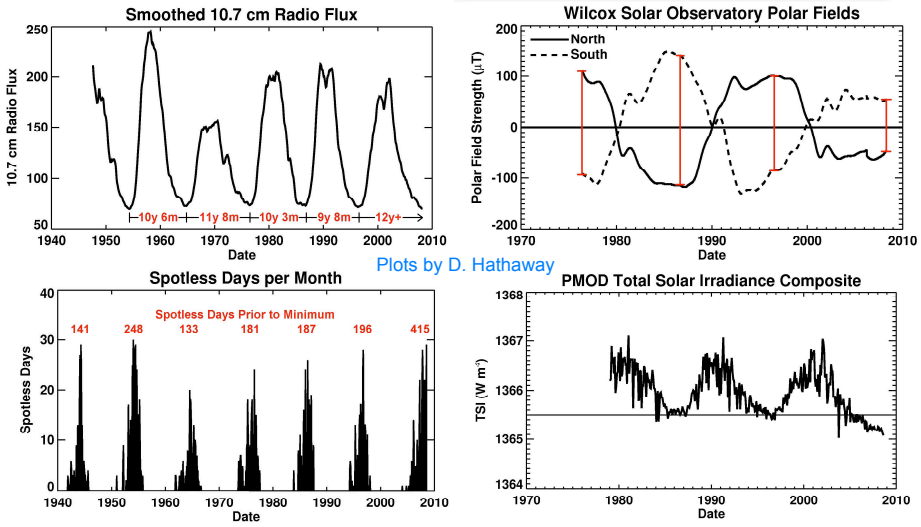
UCL

Present and next cycle



UCL

Present activity: deep minimum…

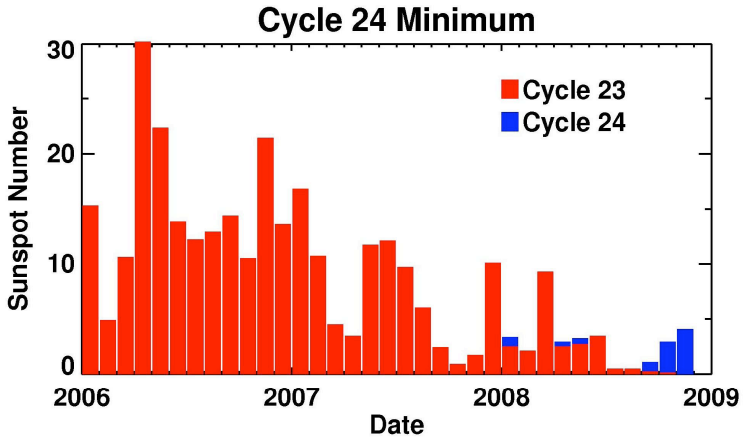


Plots by D. Hathaway

Abnormally long cycle with many spotless days, weak polar fields and decreasing solar irradiance. Will there be a cycle 24?

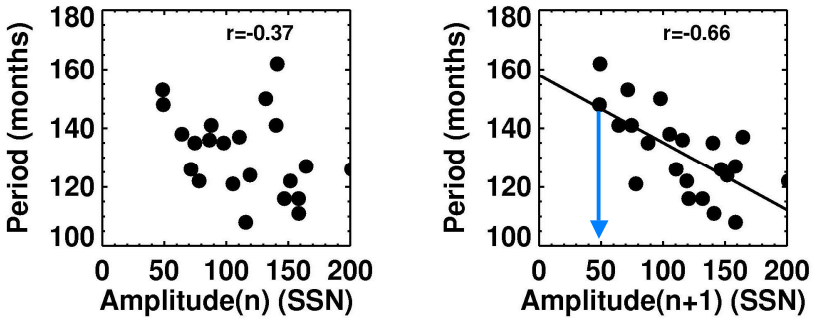
UCL

Progression of cycle 24



UCL

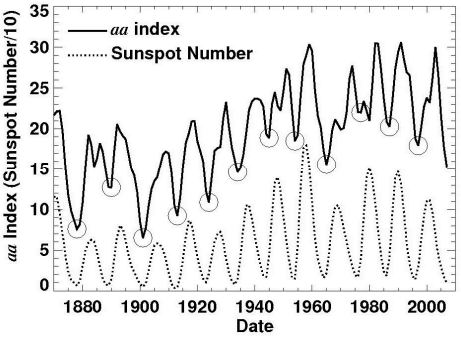
Cycle forecast

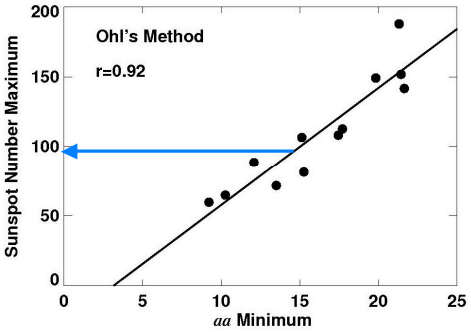


The period (length) of the cycle is related to the amplitude of the next cycle.
Big cycles tend to start early and rise rapidly leaving behind a short cycle and a high minimum.
These two characteristics suggest a small cycle 24.

UCL

Forecast: Geomagnetic precursor



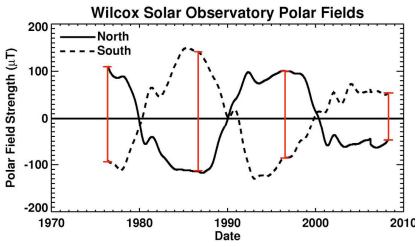


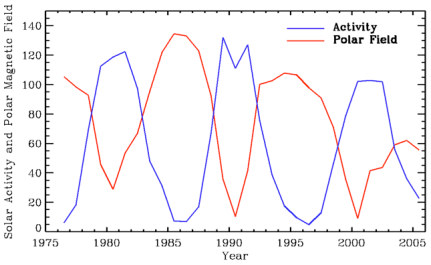
Geomagnetic activity around the time of minimum seems to give an indication of the size of the next maximum. Ohl (1966) found that the minimum in the geomagnetic index *aa* could predict the next maximum.

Note: The geomagnetic index aa at minima $\mathbf{aa_{min}}$ *is a measure of the* ***poloidal****, while the* ***sunspot number*** *at maxima is of the* ***toroidal*** *component of the solar cycle (Mayaud, 1972; Duhau, 2003)*

UCL

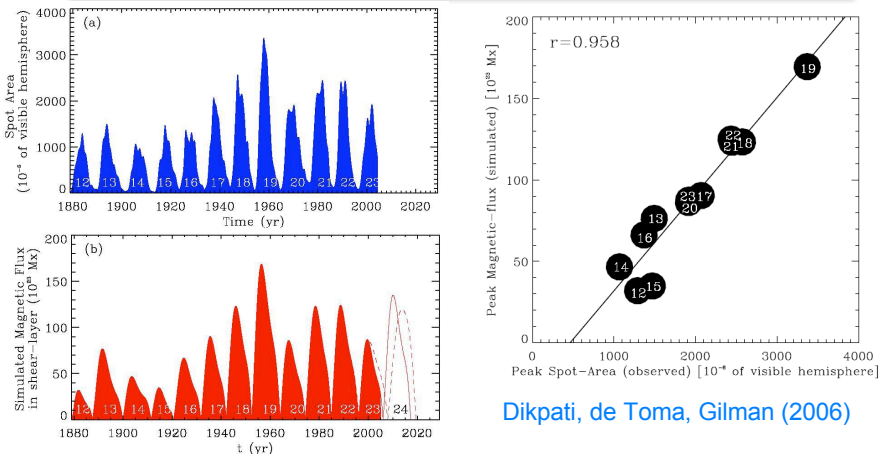
Forecast: polar field strength





Schatten et al. (1978…) have used the strength of the polar field near the time of minimum to predict the amplitude of the following maximum. Similar success to that of the geomagnetic precursors. This gives a forecast for the maximum sunspot number of cycle 24 as 75 ± 30 (Svalgaard, Cliver, and Kamide, 2005).

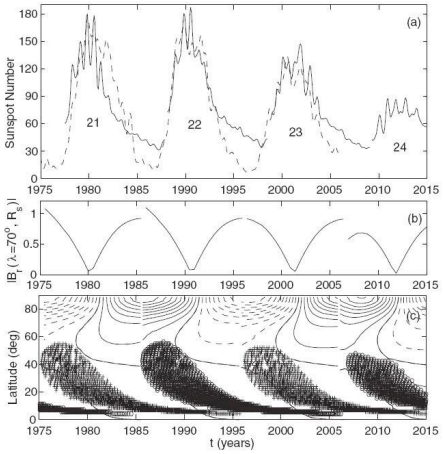
Forecast: dynamo models



Dikpati, de Toma, Gilman (2006)

Feeding sunspot areas and positions in their numerical dynamo model well reproduced the amplitudes of the last 8 cycles (RMS error < 10).
Their prediction for cycle 24 is **165 ± 15**.
Possible problem: Meridional flow is kept constant.

Forecast: dynamo models (2)



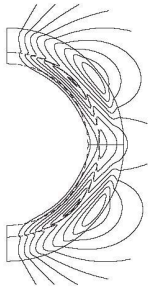


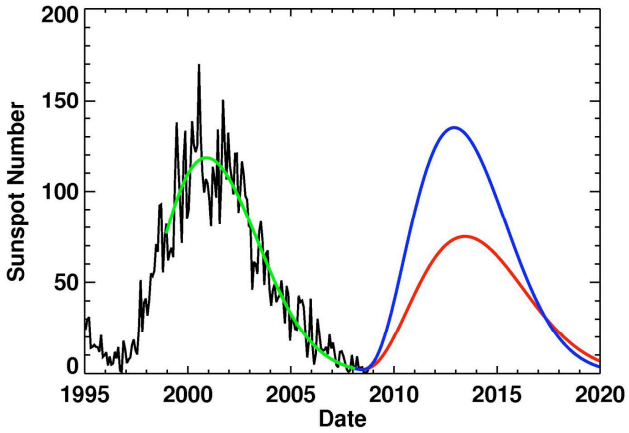
FIG. 1. A snapshot of streamlines of the poloidal field given by constant contours of $Ar\sin\theta$ just after correcting by the DM value for the poloidal field at the minimum before cycle 24. The dashed lines correspond to $r = 0.7R_\odot$ and $r = 0.8R_\odot$.

Choudhuri, Chatterjee, Jiang (2007)

Their model is dominated by diffusion. They adjust the strength of the polar field at minimum matching the observed field, but like this they "erase' memory of previous cycles. Their prediction for cycle 24 is ~75.

UCL

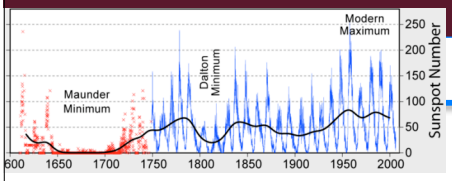
Cycle 24 predictions

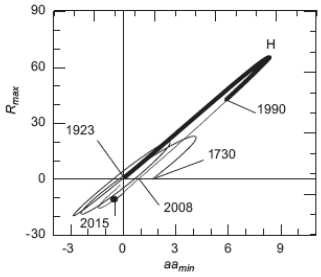


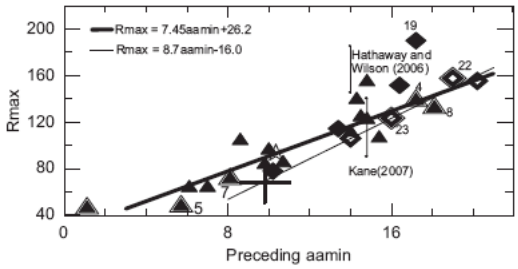
Based on a minimum in August 2008: A large cycle (R_S~135) would peak in 2012, while a small cycle (R_S~75) would peak in 2013. (Hathaway, 2009)

UCL

Cycle predictions (2)





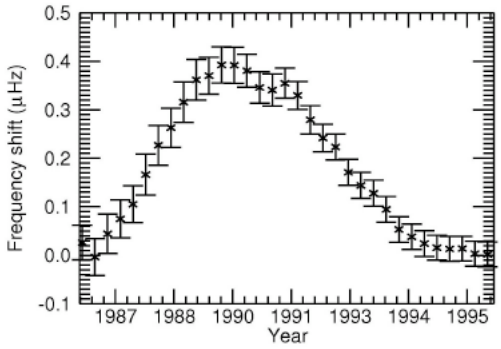


De Jager and Duhau (2009): The solar dynamo is non-linear, reflected in the cycle amplitude modulations. Decomposing the oscillations into three types + an invariant transition level they forecast cycle 24: **68 ± 17**.

They find that from 2000 the Sun is undergoing a transition from a recent *Grand Maximum* to another phase. The transition will end around the maximum of cycle 24 (~2014). Then a lower solar activity period will start, probably similar to the cycles between 1730-1923. Cycle 25 may be as low as the one ~1810, developing a short *Grand Minimum ~ Dalton minimum* lasting for 60-100 years…

UCL

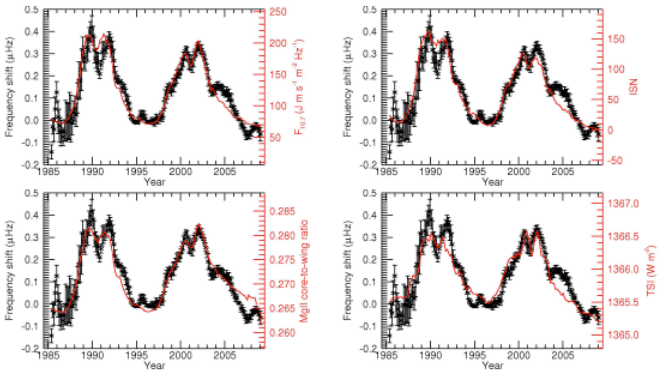
Seismic frequencies and the cycle



Solar p-mode seismic frequencies also shift with the cycle, responding to changes in the surface activity (Noyes, 1985).

UCL

Seismic frequencies and the cycle (2)



The frequency shift is confirmed using two cycles of data, and best correlated with indices which are sensitive to both the weak and strong magnetic fields.
(Chaplin et al., 2007)
But there are fine differences and the seismic frequencies are still decreasing. Have we reached the minimum? (Broomhall et al., 2009)

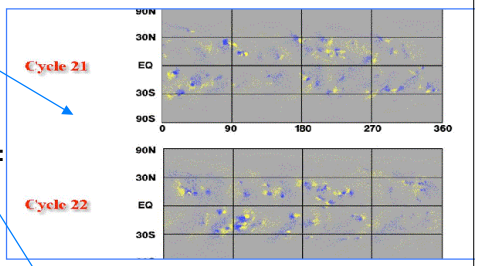
UCL

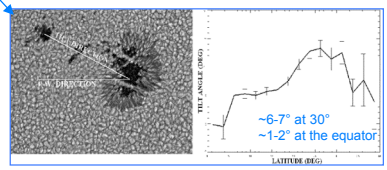
Main rules of solar activity

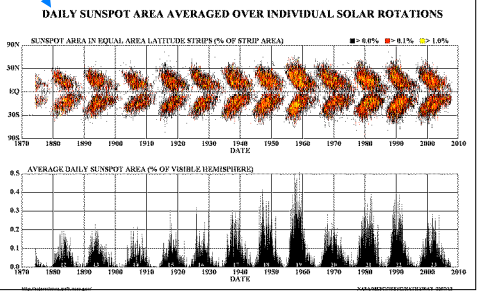
Hale(-Nicholson)'s law (Hale, 1924): active regions on opposite hemispheres have opposite leading magnetic polarities, alternating between successive sunspot cycles

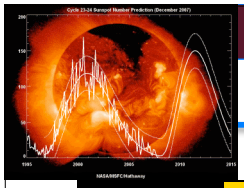
Spörer's butterfly diagram (Carrington, 1858): spot groups tend to emerge at progressively lower latitudes as a cycle progresses

Joy's law (Hale et al., 1919): the centre of gravity of the leading polarity in bipolar sunspot groups tends to lie closer to the equator than that of the following polarity; the tilt increases with latitude



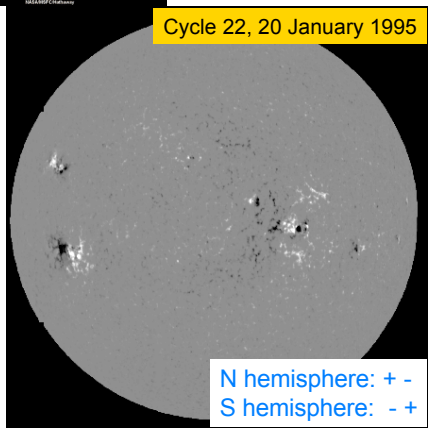


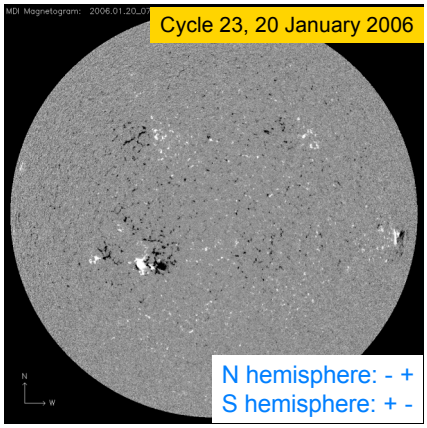




UCL

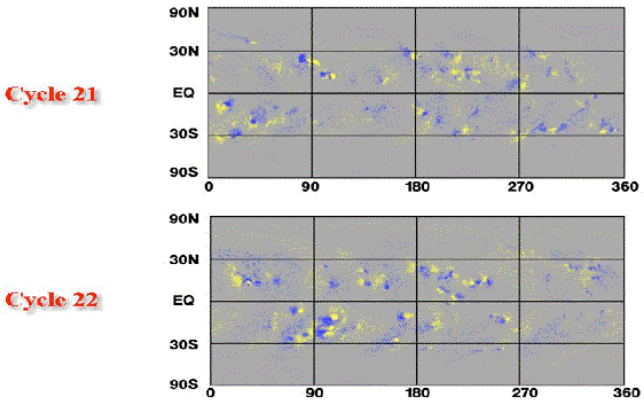
Hale's law





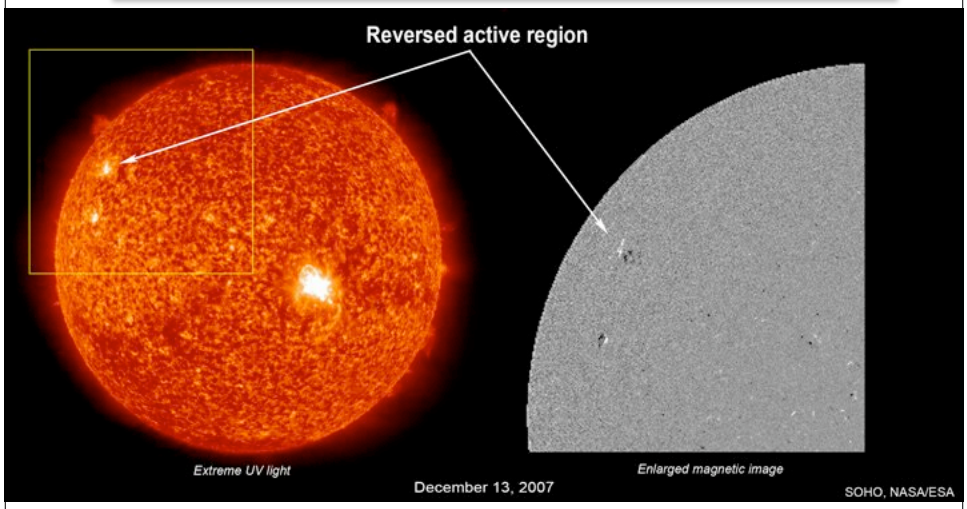
Magnetic polarities of sunspot pairs located in the northern and southern solar hemispheres are reversed; in one hemisphere the negative magnetic polarity sunspot leads the positive polarity sunspot (with respect to the westward apparent motion due to solar rotation), on the other one the reverse is true. Polarities reverse at the beginning of a new solar cycle.

Hale's polarity law



The polarity of the preceding spots in the N hemisphere is opposite to the polarity of the preceding spots in the southern hemisphere. The polarities reverse from one cycle to the next.
⇒ The magnetic cycle is, on average, 22.2 years long.

Cycle 24



The first active region of cycle 24 appeared at high solar latitude 2.5 years ago.

UCL

Hale-Nicholson rules

- The leader spot in each hemisphere are generally of one polarity, the follower spots of the opposite polarity.
- The leader polarity is opposite in opposite hemispheres.
- The magnetic axes of the bipoles are inclined, with the leader closer to the equator (Zirin named it Joy's law)
- Late in the cycle, bipoles appear at high latitudes with polarity reversed relative to lower latitudes.
- After the sunspot minimum, the prevailing leader polarity in each hemisphere is opposite to the pre-minimum sense.

UCL

Exceptions from the Hale-Nicolson rule

Many of the non-Hale oriented bipolar active regions simply belong to another cycle due to the long overlap between consecutive cycles:

- ephemeral regions appear between 40º-60 º latitude 2.5-3.5 years before the new-cycle spot groups,
- while the last regions appear close to the equator 2 years after solar minimum.

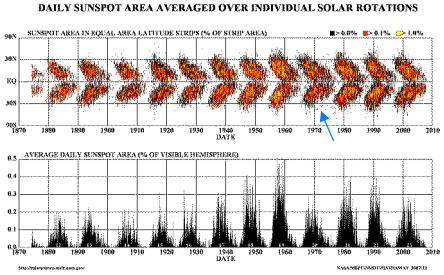
→Bipolar active regions of a given cycle continue to emerge over a total interval of 14-15 years!

Only a few percent of all bipolar active regions are reversed for their cycle and hemisphere. Estimates are 3.1% (Richardson, 1948), **4.7%** (Smith & Howard, 1968), **6.4%** (Wang & Sheeley,1989), **7.1%** (Harvey, 1993). **K. Harvey showed that this fraction increases with decreasing region size, and only 4% of large regions (A > 1000μH$_\odot$) emerge with and maintain a reversed orientation.**

Region	magnetic flux (Mx)
large	$5 \cdot 10^{21}$ - $3 \cdot 10^{22}$
small	$1 \cdot 10^{20}$ - $5 \cdot 10^{21}$
ephemeral	$3 \cdot 10^{18}$ - $1 \cdot 10^{20}$

UCL

Butterfly diagram



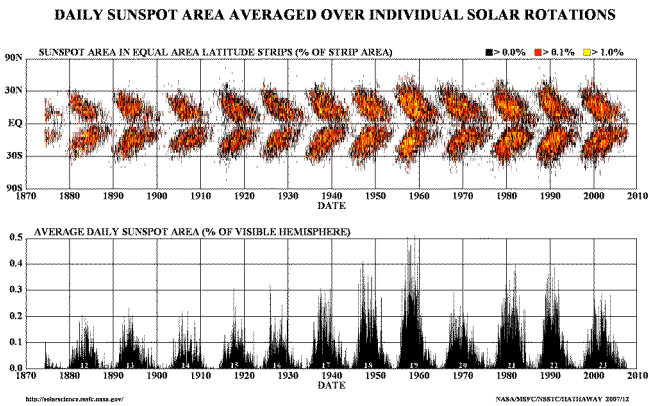
The sunspot "butterfly diagram", showing the fractional coverage of sunspots as a function of solar latitude and time (D.Hathaway).

- Sunspots are restricted to latitudinal bands ~30° wide, symmetric about the equator.
- Sunspots emerge closer and closer to the equator in the course of a cycle, peaking in coverage at about ±15° of latitude.

- Note the absence of sunspots at high latitudes (> 40°) at any time during the cycle, and the equatorward drift of the sunspot distribution as the cycle proceeds from maximum to minimum are particularly striking here.
- Note how the latitudinal distribution of sunspots is never exactly the same, and how for certain cycles (for example cycle 20, 1965---1976) there exists a pronounced North--South asymmetry in the hemispheric distributions.
- Note also how, at solar minima, spots from each new cycle begin to appear at mid-latitudes while spots from the preceding cycle can still be seen near the equator, and how sunspots are almost never observed within a few degrees in latitude of the equator.
- Sunspot maximum (1991, 1980, 1969,...) occurs about midway along each butterfly, when sunspot coverage is maximal at about 15 degrees latitude.

UCL

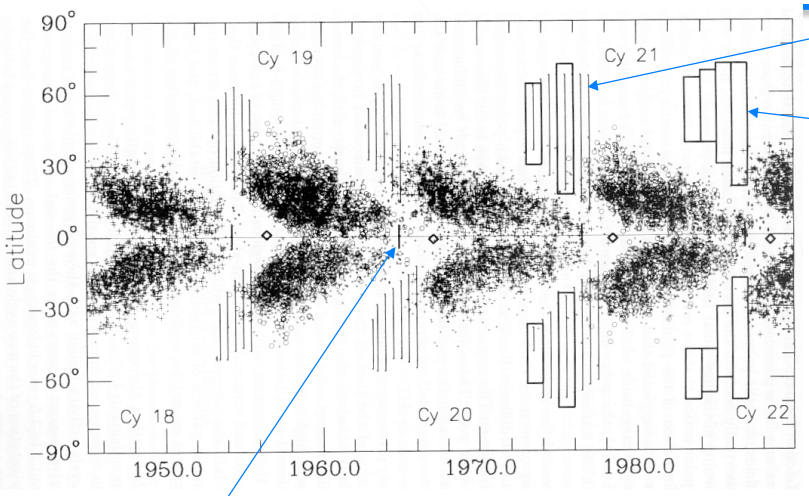
Butterfly diagram



Sunspot cycles overlap at minimum.The first sunspot group of a new cycle typically appears at high latitudes (>25°) about 18 months before minimum. The last sunspot group of an old cycle typically appears at low latitudes(<10°) about 16 months after minimum.

UCL

Butterfly diagram of sunspot regions, CaII plages and ephemeral regions from 1945 to 1991

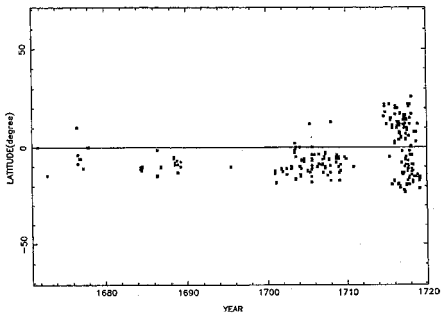


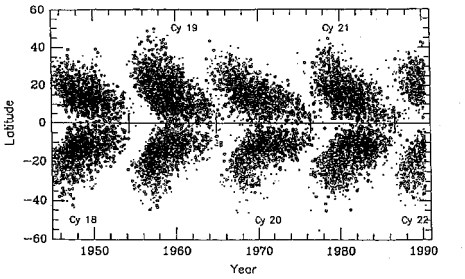
UCL

Flux emergence along the butterfly diagram

- Spot groups of all sizes emerge along the butterfly. However, the width of the distribution (spead) around the median of the butterfly diagam changes with the size of the group: the largest groups have the narrowest distribution, which is the widest for the smallest acive regions (ARs), the ephemeral regions (ERs).
- Ephemeral regions extend the wings of the butterfly diagram: new-cycle oriented ERs and small ARs are seen at high latitudes at least 3 years before the occurrence of the cycle minimum and 2.5 years before the appearance of the first new-cycle sunspot region (cycles 21/22).
- However, there is no evidence for new-cycle bipoles emerging at high latitudes before the polarity reversal. They start appearing ~ 3 years following that.

K. Harvey (1993)



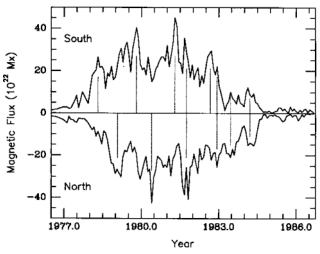


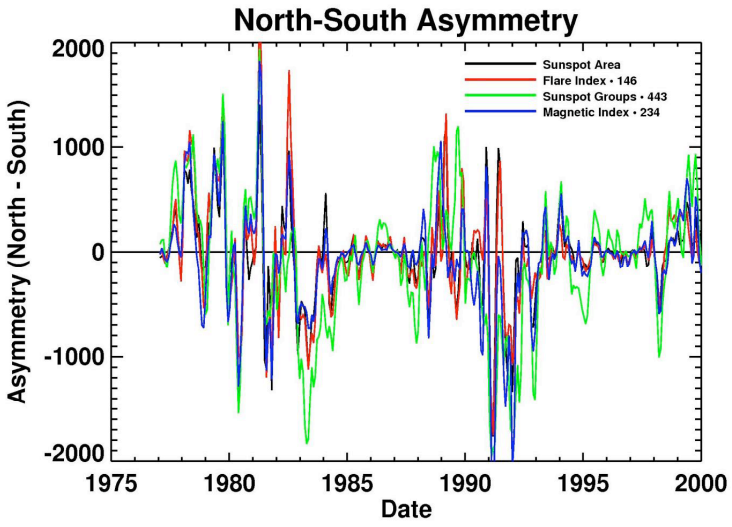
The butterfly diagram during the Maunder minimum:

- spots were seen mainly around the equator
- striking south-hemispheric dominance of activity

The study was based on ~8000 daily observations made between 1660-1719 in the Observatoire de Paris
Ribes & Nesme-Ribes (1992)

Presently, N-S asymmetry is slight (Harvey, 1993).

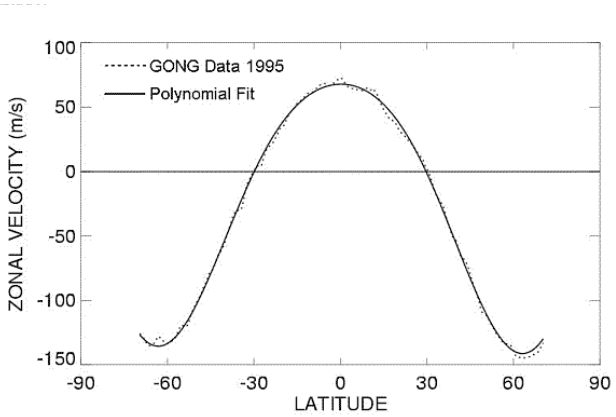




The MHD/dynamo source of the asymmetry is uncertain, but the hemispheres must be linked because they never get out of phase by much.

UCL

Torsional oscillations

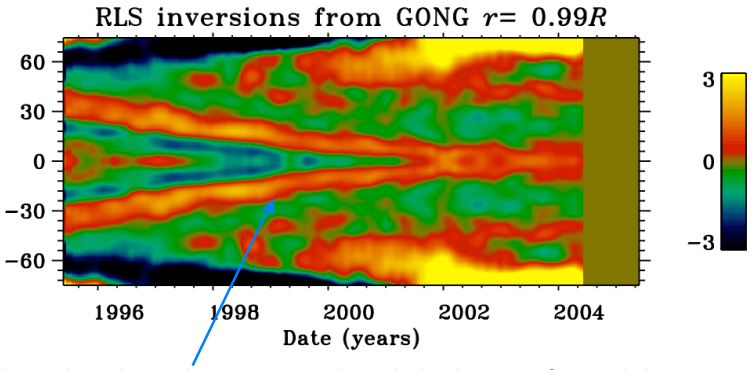


Torsional oscillations (Howard and LaBonte, 1980) are latitude bands of slightly faster and slower rotation. These bands propagate toward the equator on a solar cycle time scale giving an alternating prograde and retrograde motion at a given latitude.

UCL

Torsional oscillations - near-surface

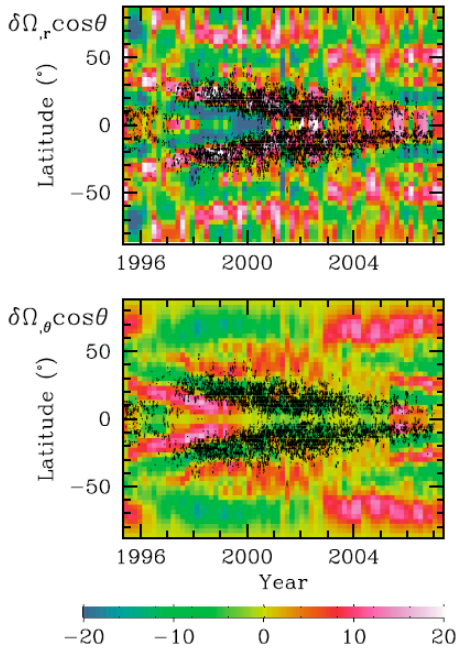
Changes of the differential rotation with the solar cycle - "torsional oscillations" (rotation speed relative to mean speed).
The variation is of the order of 1% of the rotation rate (3 nHz ~ 13 m/s).



Magnetic active regions appear at boundaries between fast and slow zones

UCL

Torsional oscillations



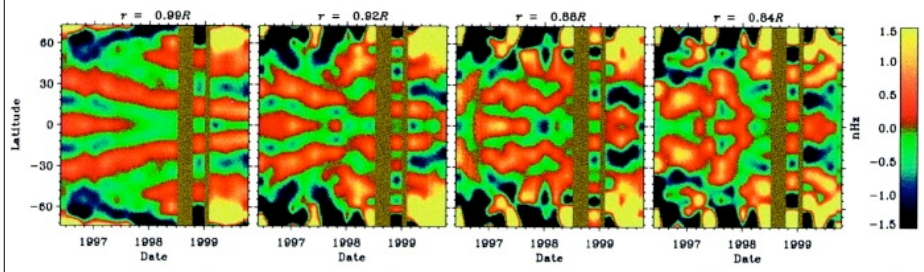
Antia et al, 2008:
Radial and latitudinal gradients of the solar rotation at r = 0.95 R_{sun}
Superposed with the butterfly diagram (sunspot locations)

Magnetic active regions appear at boundaries between fast and slow zones

The torsional oscillations may be produced by the Lorenz force acting on magnetic flux tubes moving through the convection zone (Schüssler, 1981; Yoshimura, 1981).

UCL

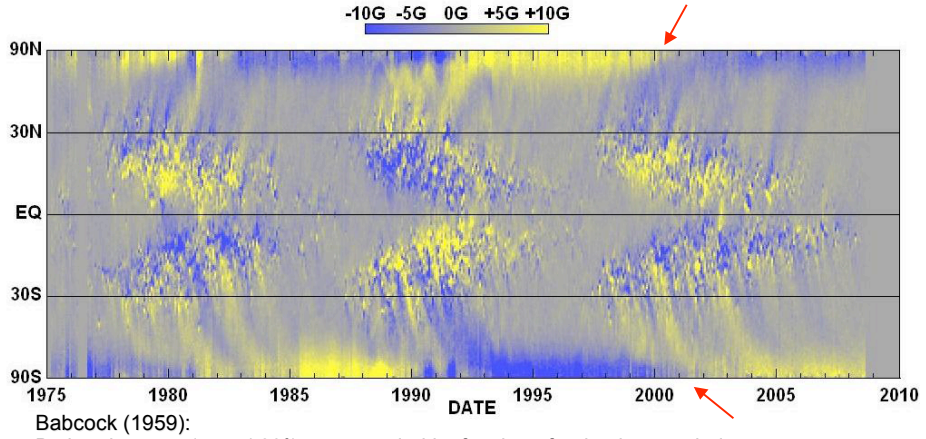
Torsional oscillations - depth dependence



The near-surface features can still be seen nearly half way through the convection zone (Howe et al., 2000; Toomre et al., 2000; Komm, Hill, and Howe, 2001)

UCL

Polar field reversal

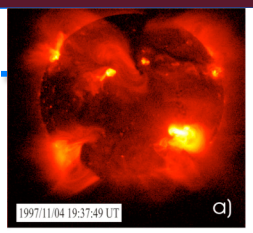


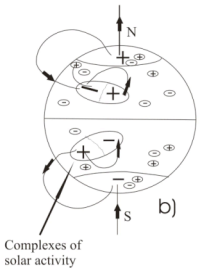
Babcock (1959):
Both polar caps (θ > +/-60°) are occupied by faculae of a dominant polarity.
They change polarity ~ 1-2 years after the solar maximum to the following polarity of the given hemisphere of the actual cycle. They are most extended and have the highest total magnetic flux just before the sunspot minimum.

Result of Joy's law + Hale's law, + meridional flow + flux cancellation across the equator.

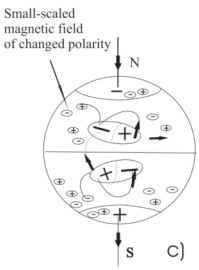
UCL

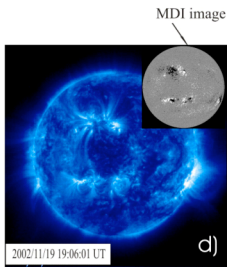
Observations





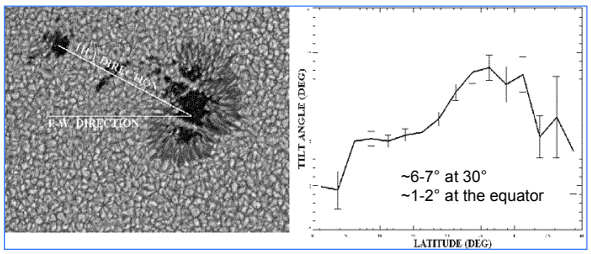
Magnetic connectivity before polarity reversal





Magnetic connectivity after polarity reversal

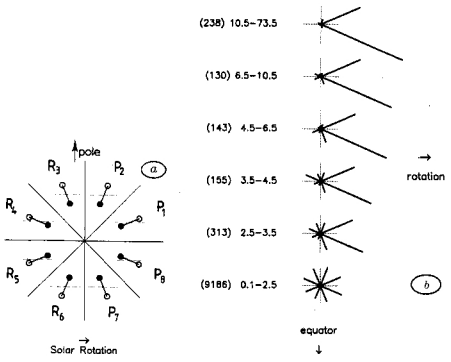
Bipole tilt - Joy's law



Observation:
Bipolar active regions show a tilt relative to the equator: the leading spot is closer to the equator than the following spot. The tilt increases with latitude. This is Joy's law (Hale et al., 1919).

Cause:
Coriolis forces acting on the rising *expanding* flux tubes leads to *clockwise rotation* of the flux tube on the *Northern hemisphere* and counter-clockwise rotation on the South.

Fluctuations of the tilt I.

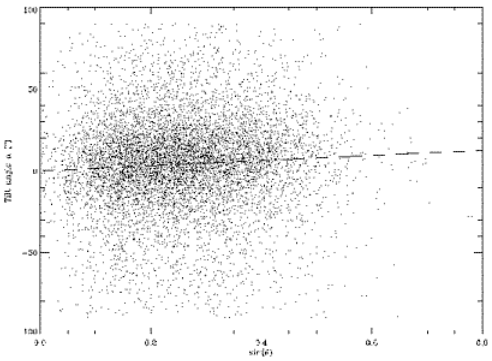


Dependence of the tilt on the size of the bipole:

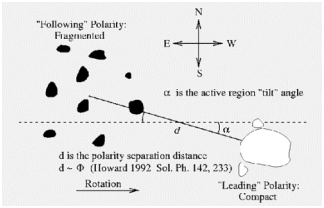
The dispersion increases with decreasing size (flux) of ARs

(Harvey, 1993)

Fluctuations of the tilt II.



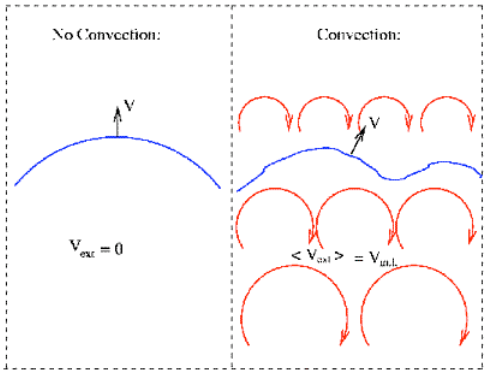
Analysis of ~ 24,000 sunspot groups shows that the tilt dispersion is not a function of latitude, but it is a function of d, with $\alpha \sim d^{-3/4}$



$\alpha \propto \phi^{1/4} \sin\Theta$

(Fisher, Fan & Howard, 1995).

The physics behind the tilt angle fluctuations

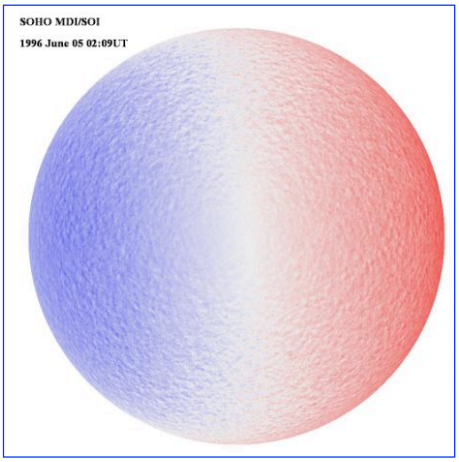


The observed variation with the size of AR suggests convective turbulence as a possible underlying cause of the tilt fluctuations.

(Longcope & Fisher, 1996)

UCL

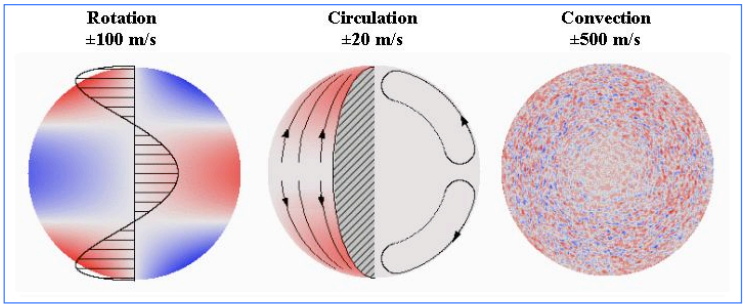
Surface flows



- Surface flows are dominated by the basic rotation of the Sun and cellular flows.
- But they also include differential rotation and an axisymmetric meridional flow.
- These flows can be measured by direct Doppler imaging, feature tracking and helioseismic inversions.
- These surface flows, and their extensions into the interior, play significant roles in the solar cycle.

UCL

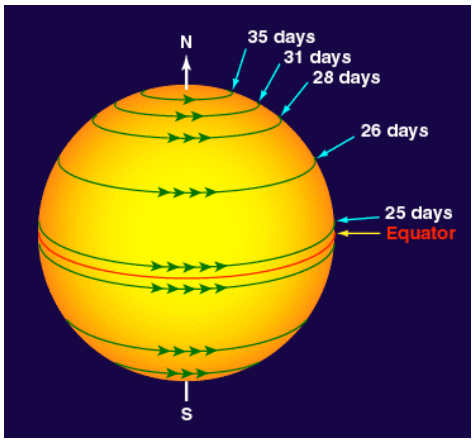
Flow components



The meridional circulation has been particularly difficult to characterize because of its weakness and masking by other velocity signals. There have been conflicting measurements since it was first measured by Beckers in 1978 (40 m/s poleward);
Others: Duvall (1979) 20 m/s poleward; Perez-Garde (1981) 20 m/s equatorward; Snodgrass (1984) 10 m/s poleward.

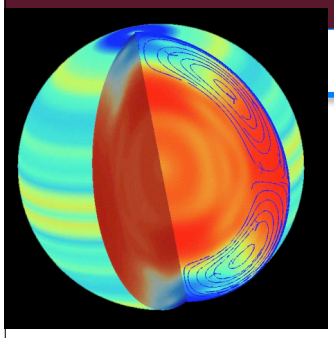
UCL

Differential rotation



Solar differential rotation: $\Omega(\phi)=A+B\sin^2\phi+C\sin^4\phi$
For magnetic features: $\Omega(\phi)=14.38-1.95\sin^2\phi-2.17\sin^4\phi$

UCL



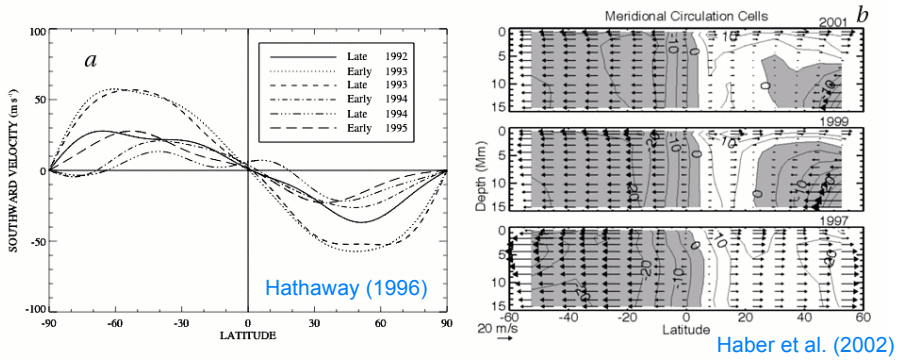
Meridional circulation

•The axisymmetric flow in the *meridional plane* is generally known as the meridional circulation.The meridional circulation in the solar envelope is much weaker than the differential rotation, making it relatively difficult to measure. Although it can in principle be probed using global helioseismology, the effect of meridional circulation on global acoustic oscillations is small and difficult to distinguish from rotational and magnetic effects. Thus, we must currently rely on surface measurements and local helioseismology.

•The meridional circulation is largely poleward from the equator at about 20 m/s and extends deep into the convection zone from the surface. Evidence is mounting that the flow is time-dependent (slower at maximum, faster at low activity).

• The equatorward return flow deep down in the convection zone has not been measured. However, it should exist (mass conservation). However, since ρ increases with depth, a 20 m/s poleward flow would be balanced by a ~ 2-4 m/s equatorward flow in the lower half of the convection zone…

Meridional circulation



Spatial and temporal variation of the meridional circulation in the surface layers of the Sun.

(a) The colatitudinal velocity in the solar photosphere obtained from Doppler measurements, averaged over longitude and time. +ve: southward flow, -ve: northward flow.

(b) The meridional flow as a function of latitude and depth inferred from ring-diagram analysis. Each inversion is averaged over a 3-month interval and results are shown for 1997, 1999, and 2001. Flow near the surface and in the southern hemisphere is generally poleward but beginning in 1998, equatorward circulation is found in the northern hemisphere at depths below.

Conclusions concerning the flows

- The meridional circulation is largely poleward from the equator at about 20 m/s and extends deep into the convection zone from the surface.
- The equatorward return flow deep down in the convection zone hasn't been measured.
- The torsional oscillation bands also display an equatorward (flow) speed and extend deep into the convection zone.
- Torsional oscillations are likely due to the effects of magnetic flux tubes threading through the convection zone and rotating at a different rate than the plasma.

Dynamo

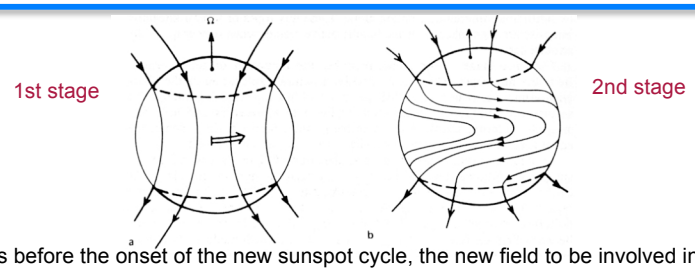
Dynamo refers to the complex of mechanisms that cause the magnetic phenomena in the Sun. Usually, it is broken down into three components:

(1) Generation of strong, large-scale fields of periodically reversing polarity
(2) The rise of these fields to the photosphere
(3) The processing, spreading across, and removal from the photosphere of magnetic flux

After decades of magnetic measurements on the Sun which included the weaker facular fields besides the strong ones, inspired by the observed time-dependent behaviour of the photospheric magnetic fields, in 1961 Babcock put forward a conceptual model, which describes the 22-year cycle in five stages.

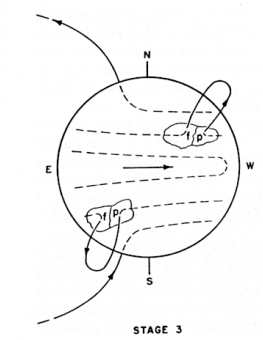
Babcock's model

:



1st stage:
About 3 years before the onset of the new sunspot cycle, the new field to be involved in the new cycle is approximated by a dipole field symmetric about the rotation axis.
That poloidal field is identified with observed field in the polar caps.
Low-latitude regions are residuals of the preceding cycle.
All the internal poloidal field lines are arbitrarily taken to lie in a relatively thin sheet at depth in the order of 0.1 $R_{\odot}$ that extends in latitude from -55^{o} to +55^{o} having flux of 8x10^{21} Mx.

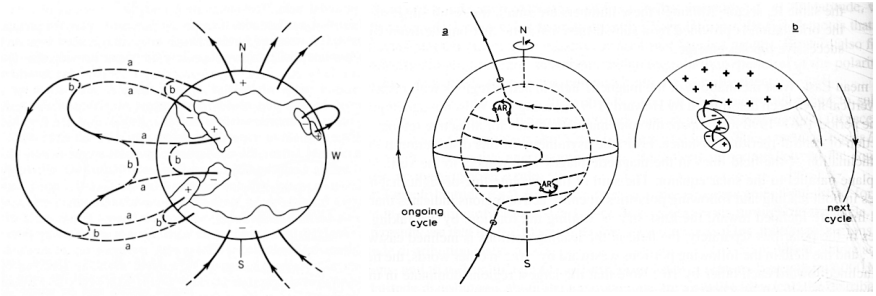
2nd stage:
The originally poloidal field is pulled into a helical spiral in the activity belts, with the result that the field becomes nearly toroidal and gets amplified - after 3 years the equator will have gained in its rotation about 5.6 turns on the latitude circles +/- 55^{o}, which gives the same number of turns for the field lines on both hemispheres (0.0001→0.03T). Further amplification was reached by **twisting** by the roller bearing effect of the radial differential rotation shear: ***α-effect*** (→0.3T)



Third stage: Formation of active regions

- Provides a natural explanation of the Hale-Nicolson polarity law. Each Ω-shaped loop erupting through the surface will produce a bipolar active region with preceding and following magnetic polarity, and p and f polarities switch at the equator.
- Explains Spörer's law: given the $\sin^2\phi$ (ϕ=latitude) term in the solar differential rotation, the intensification proceeds most rapidly at ϕ=+/-30°, so the 1st ARs of the cycle expected to erupt there, and only later at lower latitudes.
  It takes 8 years between the 1st spots at ϕ=+/-30° and the last low-latitude spots to appear using the differential rotation curve.

Solar differential rotation: $\Omega(\phi)=A+B\sin^2\phi+C\sin^4\phi$
For magnetic features: $\Omega(\phi)=14.38-1.95\sin^2\phi-2.17\sin^4\phi$



Fourth stage:
Describes the neutralization and subsequent reversal of the general poloidal field as a result of the systematic inclination of active regions (i.e. that the following polarity tends to lie at higher latitude than the leading polarity). When opposite polarities are brought together, equal amount of opposite flux cancel.
~99% of AR flux cancels against remnant flux from neighbour ARs, less than 1% of following polarity makes it to the nearest pole, first neutralizing the existing field and then replacing them by flux of opposite polarity. The same fraction of leading polarities of the two hemispheres cancel in the equatorial strip.

UCL

Fifth stage:
Babcock assumed that ~11 years after the beginning of the 1st stage the polar fields correspond to a purely poloidal field opposite to that during the 1st phase. From here on, the dynamo process has been suggested to continue for the second half of the 22-yr activity cycle, now with all polarities reversed with respect to the first 11-yr.

In Babcock's model, the poleward migration of the f polarity towards the nearest pole was included simply as an observed fact.

Leighton (1969) interpreted the mean flux transport as the combined effect of
- the dispersal of magnetic elements by a random walk process
- and (like Babcock) the asymmetry in the flux emergence (tilt - Joy's law)

Leighton included this flux transport in a quantitative, closed ***kinematic model*** for the solar cycle, dealing with zonal and radial averages of the magnetic field, which are assumed to vary only with latitude and time.
→ ***Babcock-Leighton semiempirical kinematic model***

UCL

Limitations of the B-L model

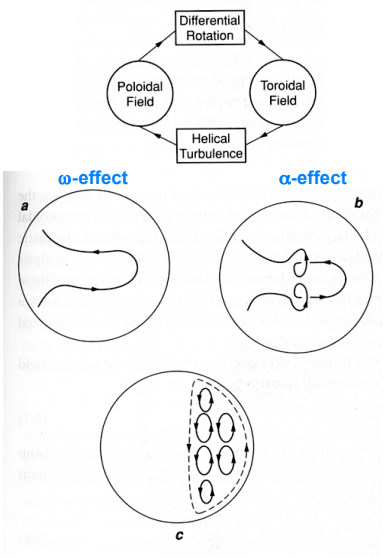
• Heuristic and semi-quantitative

• Kinematical - velocity field v is assumed, while the dynamics of the magnetic phenomena are much less understood

• In the original Babcock model the internal poloidal field lines were arbitrarily taken to lie in a relatively thin sheet at a depth of 0.1 $R_\odot$, while now it is commonly accepted that the toroidal flux cannot reside in the convection zone proper, but it is stored in the thin boundary layer below the convection zone called the tachocline.

• The intensification of the field is produced by both latitudinal and *radial* gradients in the Sun's angular rotation rate. The latter impart a further twist to the flux tubes, adjusting the α effect to the required level in order to reproduce Spörer's law. However, the radial gradient in the differential rotation (in the convection zone) is not confirmed by helioseismology - however, there is radial gradient in the tachocline.

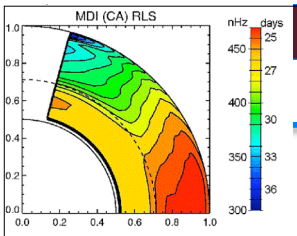
Recent models



Since then, dynamo models have aimed fully ***dynamical solutions*** of the induction equation together with the coupled mass, momentum and energy relations for the plasma
→ dynamo equation.

All dynamo models rely on the differential rotation to pull the field in the horizontal direction (ω-effect).
The main problem of the solar dynamo is how to generate the properly cyclic poloidal field.

Parker (1955): rising plasma blobs (in the CZ) expand laterally → they rotate due to the Coriolis force → helical turbulence in convective motions →rising magnetic elements carry a poloidal field component correct for the ***next cycle.***



With ***helioseismology*** establishing the strong differential rotation shear at the base of the convection zone in the ***tachocline***, there is little doubt that the ***dynamo operates there***, and the magnetic field produced there has a magnitude of the order of 10 T.

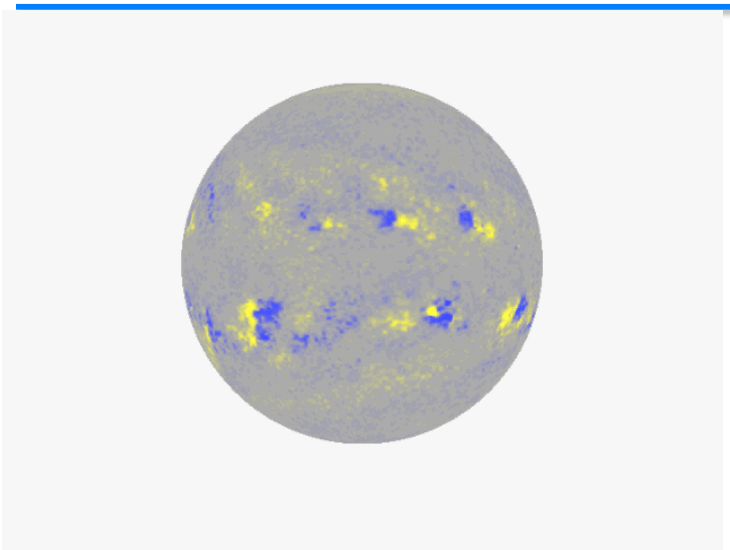
However, such strong field can not be twisted by helical turbulence, so for the ***generation of the poloidal field*** the BL idea is revived - i.e. that it is produced by the *decay of bipolar regions on the surface*.

It is the ***meridional circulation*** which will carry this field poleward and then down to the base of the convection zone, where it is stretched by the differential rotation to produce a strong toroidal field.

→ ***advective dynamo model*** (Choudhuri et al., 1995).

Though the details are far from being clear, it seems that ***solar magnetic fields are generated and maintained by the dynamo process***.

Movie of two full magnetic cycles

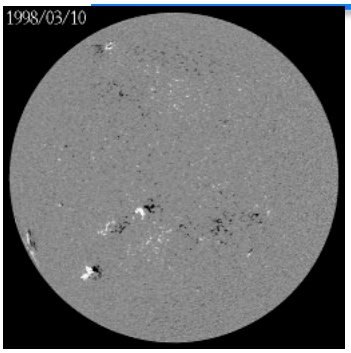


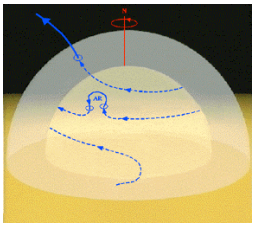
Movie by
D. Hathaway

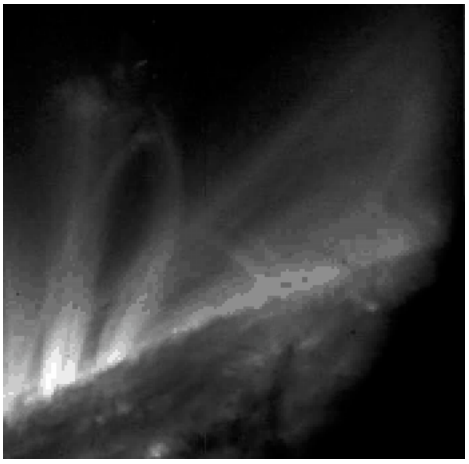
The movie shows well the 11-year cycle, AR migration, Joy's law, Hale's law, polar field reversal, and the effects of differential rotation and meridional flow.

UCL

Magnetic flux emergence on the Sun







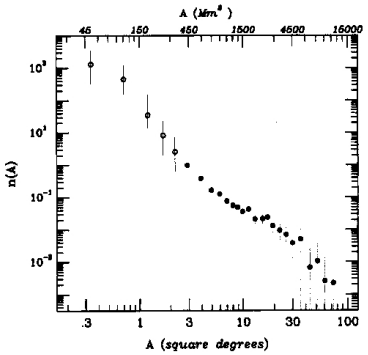
UCL

Characteristics of flux emergence

- Classical picture of flux emergence: (fragmented) Ω-loop
- Main characteristics of bipolar active regions (ARs)
- Multi-wavelength detailed picture of flux emergence
- Asymmetry in the Ω-loop – eastward tilt
  ephemeral regions – E- W asymmetry – westward tilt?
- Signatures of twist in emerging flux tubes
- Emergence of flux tubes deformed by deep CZ vortices

Size distribution of spots

Region	magnetic flux (Mx), one polarity	lifetime
large (with sunspots)	$5 \cdot 10^{21}$ - $3 \cdot 10^{22}$	weeks-months
small (no spots, maybe pores)	$1 \cdot 10^{20}$ - $5 \cdot 10^{21}$	days-weeks
ephemeral	$3 \cdot 10^{18}$ - $1 \cdot 10^{20}$	hours (mean: 4.4)

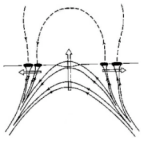


Cycle-averaged size distribution of 978 bipolar regions at max. development + 9492 ephemeral regions (ERs) Zwaan & Harvey, 1994

Scenario of magnetic flux emergence

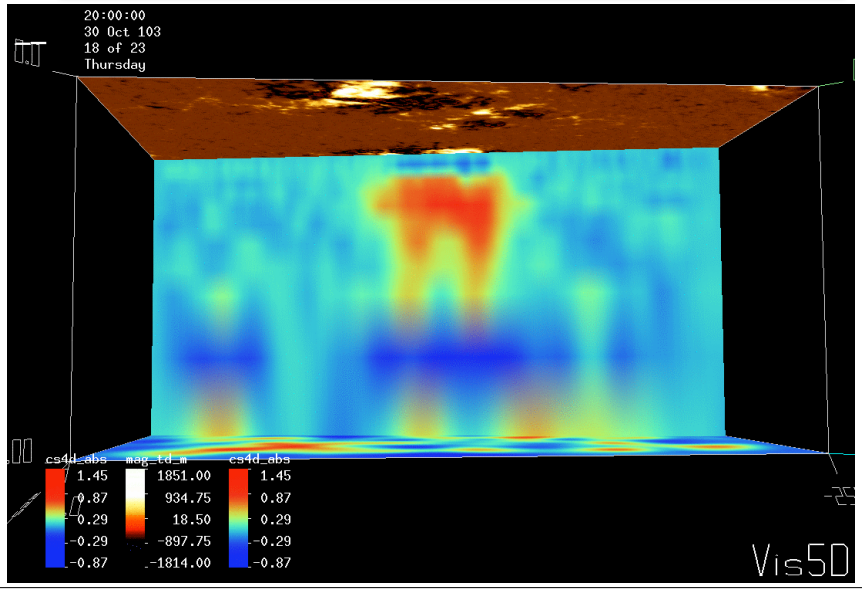
Formation of bipolar active regions:

- appearance of a small, bipolar bright plage in Ca II H&K line (Fox, 1908)
- arch filament system in Hα connects faculae of opposite polarity (Bruzek, 1967) with ascending loop top ($v \leq 10$ km/s) and draining material along the legs ($v \leq 50$ km/s), lifetime of individual fibrils ≈ 20 min
- hot (bright) EUV and X-ray loops appear above the AFS loops
- dark photospheric lanes (alignment of granulation) parallel to the overlying AFS with lifetime of ≈ 10 min, $v \approx 3$ km/s, **B** ≈ 600 Gauss
- opposite polarities move apart ($v \leq 2$ km/s – first half hour
  $v \approx 1.3$-0.7 km/s next six hours) (Harvey & Martin, 1973)
- sunspots form by coalescence of pores/smaller spots
- bipole orientation may be arbitrary initially, but in about 1-4 days it becomes correct – nearly parallel to the equator, with the p polarity closest to it (Joy's law).
- this sequence of events is pictured as the emergence of Ω-loops.



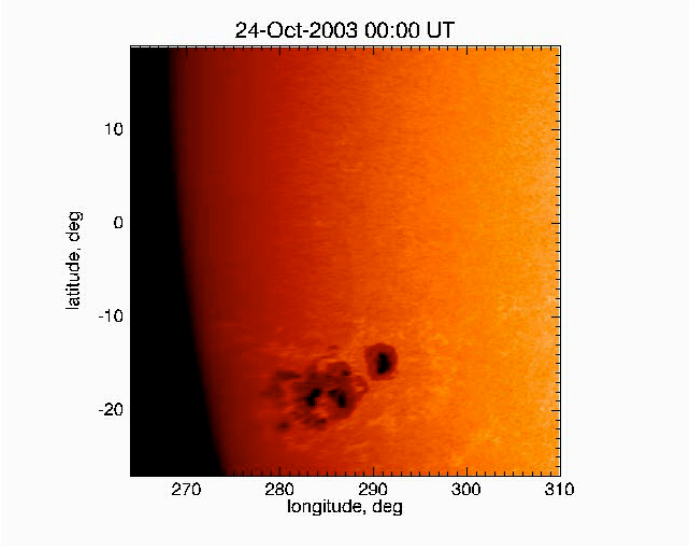
UCL

Flux emergence seen below the surface



UCL

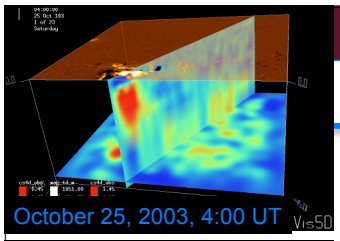
Sunspot group formation



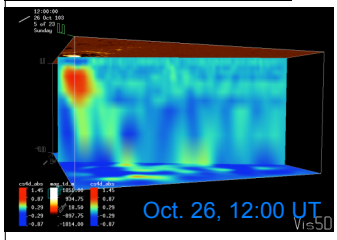
Evolution of AR 10486-488: October 24 – November 2, 2003

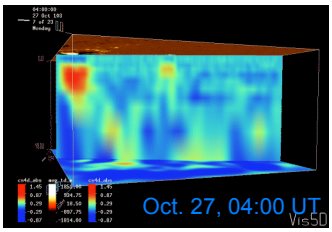
UCL

Flux emergence seen sub-surface



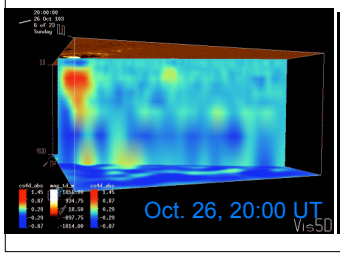
Sound-speed maps and magnetograms of AR 10486 (depth of the lower panel: 45 Mm)

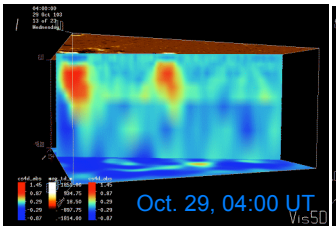


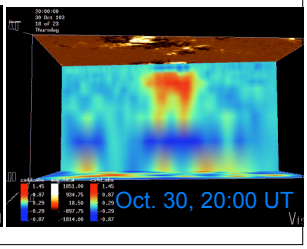


Emerging magnetic field seen in time-distance data

A. Kosovichev

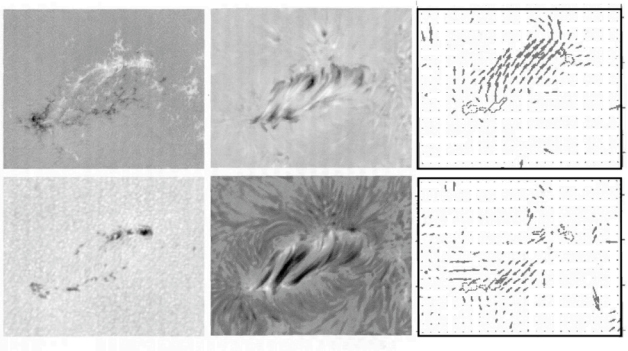






UCL

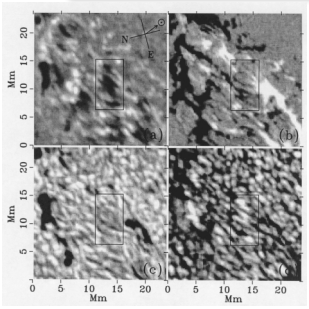
A well-observed flux emergence

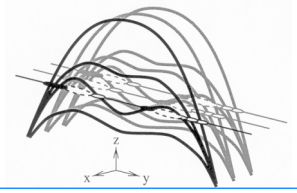


- elliptical shape outlined by pores (long axis: 55 Mm) separation rate of pores of opposite polarities: 0.73 ±0.07 km/s (v_p=0.46 ±0.05 v_f=-0.28 ±0.05)
- counter-streaming of facular points of opposite polarities (v_p=0.61 ±0.04 v_f=-0.24 ±0.05), highest shear velocities are in the middle of the EFR

Strous 1994; Strous et al, 1996; Strous & Zwaan 1999 :
AR 5617 on 29-July-89: 7.5 hours after the emergence began, observed an EFR for 1.5 hour with the SOUP tunable filter at the Swedish VST on La Palma

Fine structure of flux emergence

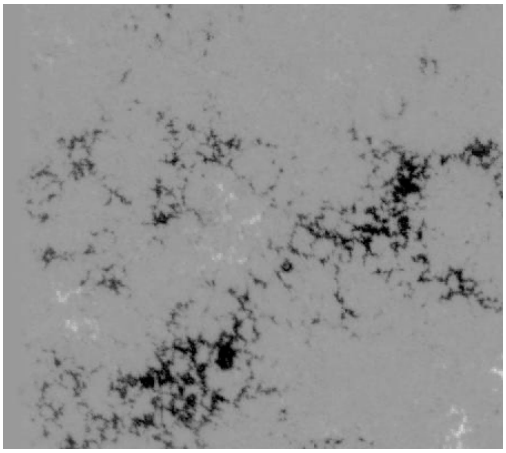




Shows such a complexity of detail that is not described by the simple emergence model of an Ω-loop.

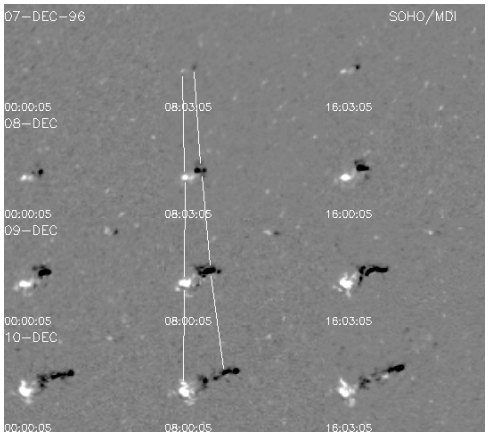
- Small-scale emergence events are characterised by transient darkening of lanes (2 Mm, 10 min) in the continuum → upflows, weak or no magnetic field followed by bright facular grains at their end(s) grains → downflows, magnetic field concentrations
- There is not a single magnetic inversion line, both polarities are found all over the AR.
- Flux emergence happens recurrently in a number of locations all over the AR, with a $\lambda = 8$ Mm.
- Facular motion, unipolar facular alignments, emergence locations, flux emergence events, subsequent footpoint motion, separation, Hα loops all line up in the same direction.
- **The emerging flux tube is frayed in two systems:**
  (i) in vertical stacks arranged in slightly curved, nearly parallel sheets
  (ii) many flux tubes emerge in multiple locations (undulatory flux tubes)

A Hinode/SOT observation



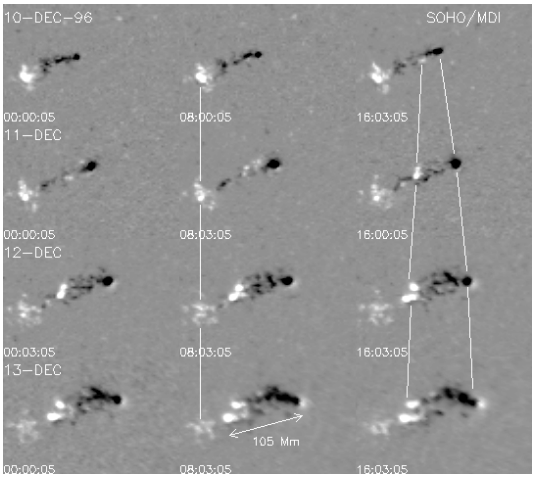
The see-serpent like pattern results from fact that the emerging flux rope has great difficulties in crossing the photosphere due to fast-changing physical conditions (e.g. Magara, 2004; Manchester et al., 2004) leading to its fragmentation. Flux emergence must involve many episodes of magnetic reconnection to succeed (Pariat et al., 2004)

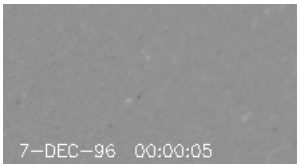
Emergence of a bipolar active region



Note the main features:
- as the p & f spots separate, p moves faster westward than the f spots eastward
- the p spots form an elongated pattern of negative polarity in the direction of their motion → asymmetric inversion line

Emergence of a simple (?) bipolar AR

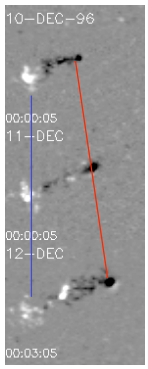




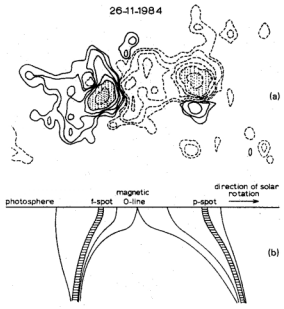
Large active regions are formed by **several bipoles** that surface separately but in **close proximity**, in **rapid succession** within a few days (Schrijver & Zwaan, 2000).

Asymmetry

in motion of opposite polarities

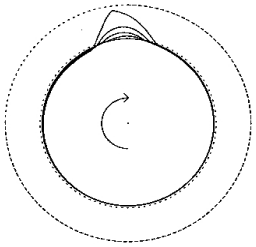


magnetic field distribution



van Driel-Gesztelyi & Petrovay (1990)

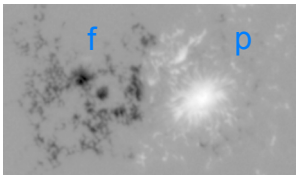
model viewed from the pole

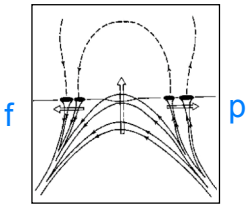


Caligari et al. (1995)

Emerging flux tubes are systematically inclined eastward, trailing the solar rotation
Cause: effect of conservation of angular momentum during their ascend in the convection zone

Asymmetry in stability

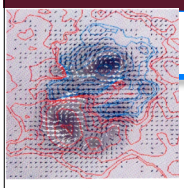


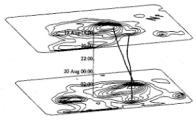


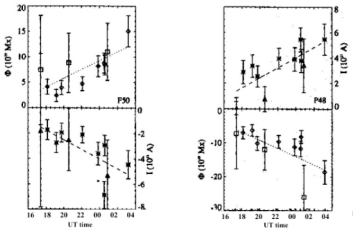
Observation: preceding spots are larger and more stable than following spots

Cause: plasma flows from preceding towards following spots induced by the Coriolis force while ascending in the convection zone. The flows increase the gas pressure in the f branch. (Fan et al. 1993; Fan & Fisher, 1996; Sigwarth et al., 1998)

Emergence of twisted flux







Leka et al, 1996 -- AR 7260 in August 1992
Evidence:

- Hα and X-ray structures associated with the emerging bipoles do not agree with potential magnetic extrapolations
- proper motions of spots indicate the emergence of deformed (kink?) emerging flux tube geometry

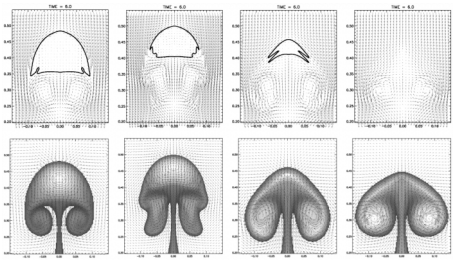
- the bipoles are co-spatial with significant vertical electric currents
- all the above signatures imply the same sense of twist
- **currents increase together with the emerging flux**

Wang & Abramenko (2000) found similar results from the analysis of AR 7321.
Other authors (e.g. Ishii et al. 2000) use Hα morphology and sunspot motions to argue that flux emerges twisted.

Since then **photospheric helicity flux** provided decisive supporting evidence (Moon 2001; Moon et al., 2002, Nindos & Zhang, 2003; Chae et al., 2004, Yamamoto et al., 2005…)

Twist is needed to keep the flux' integrity

Pitch angle: 13.9^0 7^0 2.5^0 0^0



Model:

- untwisted flux is destroyed by vortices which form in its wake in the CZ (Longcope, 1996)

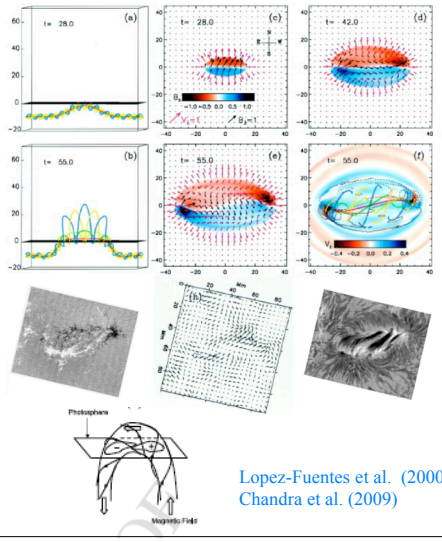
- twist can conserve the integrity of the rising flux tube
  ⇒ all major magnetic flux which has crossed the convection zone and emerge ***must be twisted.***

(Emonet & Moreno-Insertis, 1998)

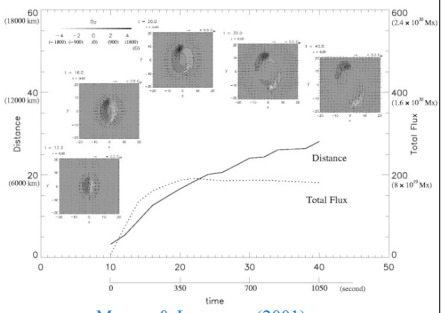
UCL

Even Strous' AR was twisted…

Fan (2001)



Lopez-Fuentes et al. (2000)
Chandra et al. (2009)

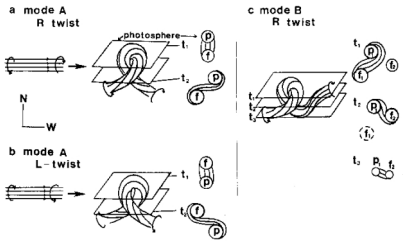


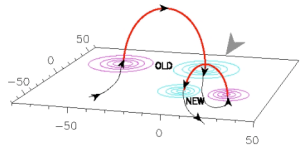
Magara & Longcope (2001)

Simulations of emergence of twisted Ω-loops provide pictures amazingly similar to those observed by Strous (1994)…
AR 5617 showed all the main characteristics of emerging ARs: it was both **asymmetric** (inclined flux) and **twisted**.

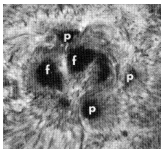
UCL

If the flux is even more twisted…

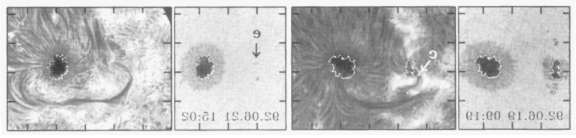




then we can get a kinked Ω-loop, which may form a « stitch » (Pevtsov & Longcope, 1998) ...



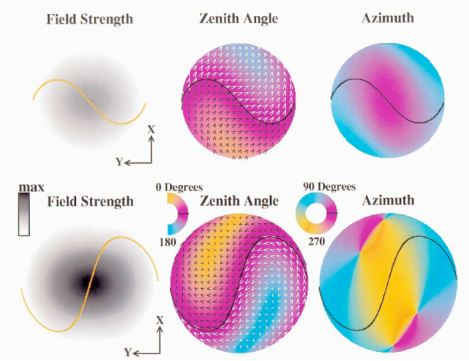
or form « knotted » δ-spots (Tanaka, 1991), which are born and die locked together (Zirin & Liggett, 1987) …



and sometimes can even form a flux « ball »…

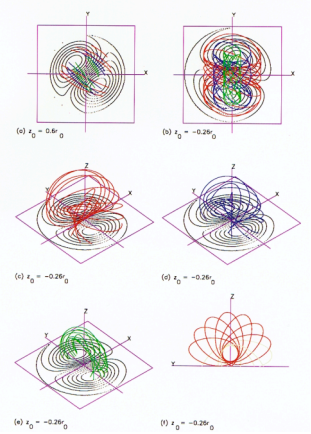
UCL

Emergence of a flux « ball »



Distribution of the vector magnetic field in the photosphere (Lites, Low et al. 1995) .

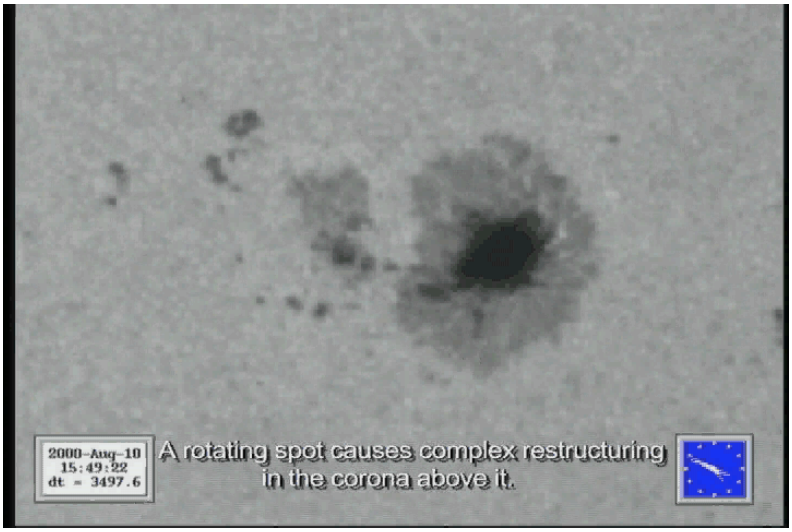
Is it the result of a disconnected knotted flux tube, which went through subsequent relaxation?



3-D field line geometry

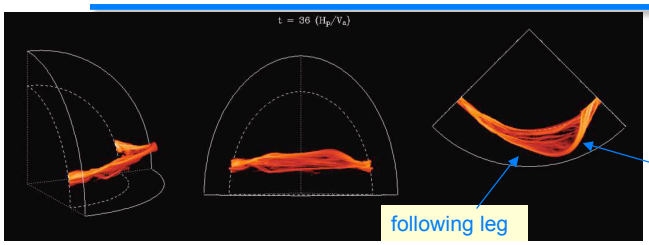
UCL

Strong twist: rotating sunspots



UCL

Asymmetry & twist revisited



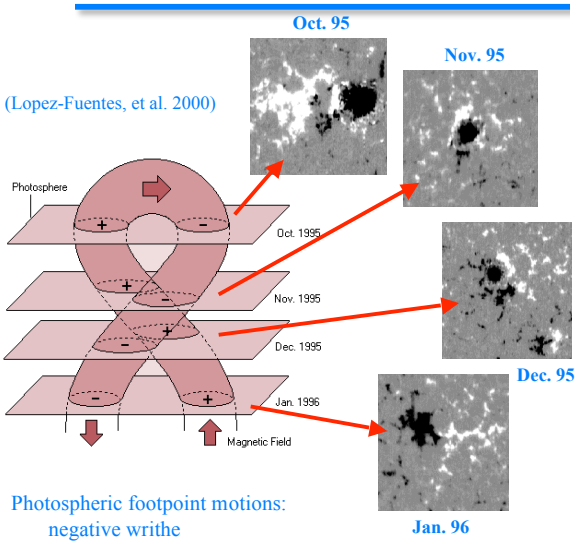
Fan (2008)

Fan (2008): ***Asymmetry in the shape of the flux tube is opposite then previously thought.***
Asymmetry is due to asymmetric stretching of the rising flux tube by the Coriolis force
⇒ asymmetric field strength in leading and following legs (leading field gets stronger)
⇒ leading leg becomes more buoyant, leading to an asymmetry in the shape of the loop, which is opposite to the asymmetry found in earlier simulations in thin flux tube approximation.
⇒ the asymmetry in sunspot motions must then be due to the faster emergence of the more buoyant leading leg (van Driel-Gesztelyi and Culhane, 2009).

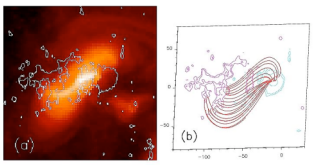
Fan (2008): ***Twist in the flux tube is lower then previously thought.***
Twist-induced tilt at the flux tube's apex is opposite to that induced by the Coriolis force - the latter is the observed tilt (Joy's law)
⇒ twist must be lower than the one needed for a cohesive rise of the flux tube
⇒ 50 % of the initial flux remain in the Ω-loop by the time it reaches the surface.

UCL

Flux tube deformed by sub-photospheric vortices

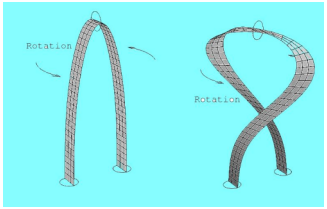


Photospheric footpoint motions: negative writhe

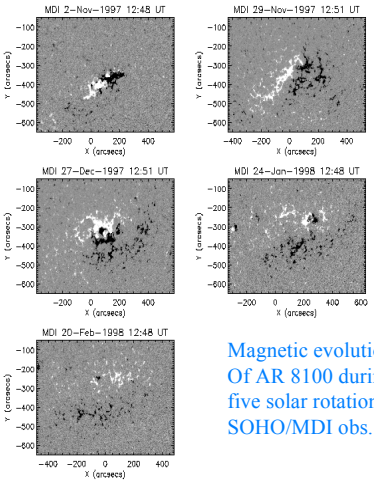


Extrapolations: positive twist

The deformation of the flux tube can not be due to kink instability!



The AR starts as a normal Ω loop at the overshoot region and, as it travels through the CZ, it is deformed by deep-rooted motions having a rotational component.

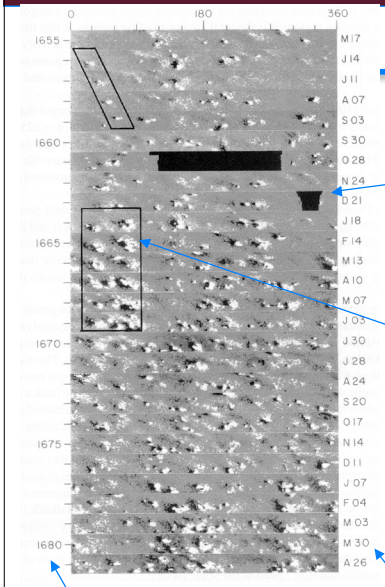


Magnetic evolution Of AR 8100 during five solar rotations, SOHO/MDI obs.

A Hale-oriented region which turned to become non-Hale…

Another example of the emergence of similarly distorted flux tube (Green et al. 2002).

Such deformed flux tubes are rare, but not unique, so the role of deep-rooted vortices seem to be of some importance…



Carrington longitude

Carrington rotation No.

Date of the beginning of the rotation

Large-scale patterns in flux emergence

Nests of activity between May 1977-May 1979 after Gaizauskas et al. (1983).

30°-wide latitudinal strips of magnetic synoptic maps

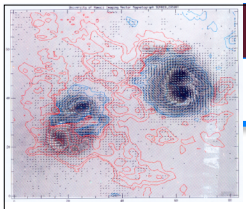
Clustering of sunspot appearances was first noticed by Cassini (in 1729!).

In this complex 29 major AR emerged, but the total unsigned magnetic flux remained constant at 1.3 x 10^{22} Mx.

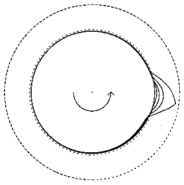
Harvey (1993): bipolar ARs tend to emerge within existing bipoles, with a separation of 4-5 days. Harvey & Zwaan (1993) found a 22-fold higher emergence rate in existing regions than elsewhere.

More on nests…

- Nearly 50% of all bipolar regions are members of an active nest.
- In a nest, members are similarly oriented and have similar complexity.
- Most of the magnetic flux seems to disappear within the boundary of nests (nearly equal emergence/cancellation rate).
- Nests have higher-than-average activity due to the interaction of bipoles.
- After flux emergence stops, the magnetic field starts spreading.
- Statistically, there are no nests along the same meridian → no active longitudes!



Summary on flux emergence



Emerging flux tubes are inclined to the vertical, twisted, distorted, frayed and fragmented.

- In spite of being twisted, emerging flux tubes get fragmented and frayed
- ARs show a latitude-independent dispersion of tilt – effect of turbulence

top of CZ physics

- Rising flux tubes can be distorted severely by deep vortices in the CZ

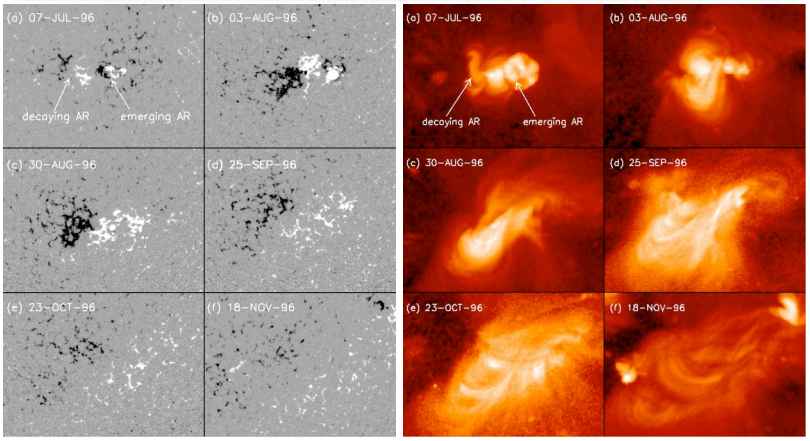
both

- Flux tubes are inclined to the vertical – in which sense?
- Flux tubes undergo Coriolis rotation during their rise (Joy's law))
- The flux emerges with inherent twist.
- Flux emergence has a clustering tendency

CZ and tachocline physics

UCL

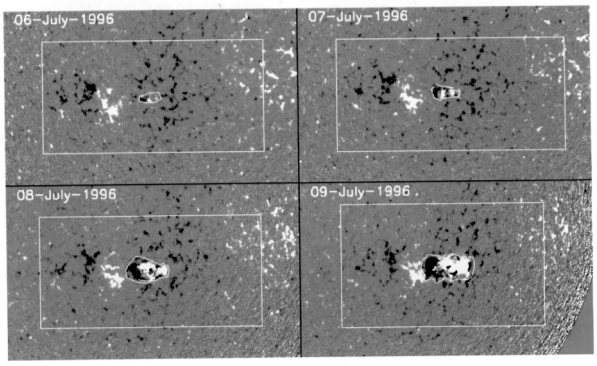
Long-term evolution of an AR – emergence & decay



SOHO/MDI Yohkoh/SXT

UCL

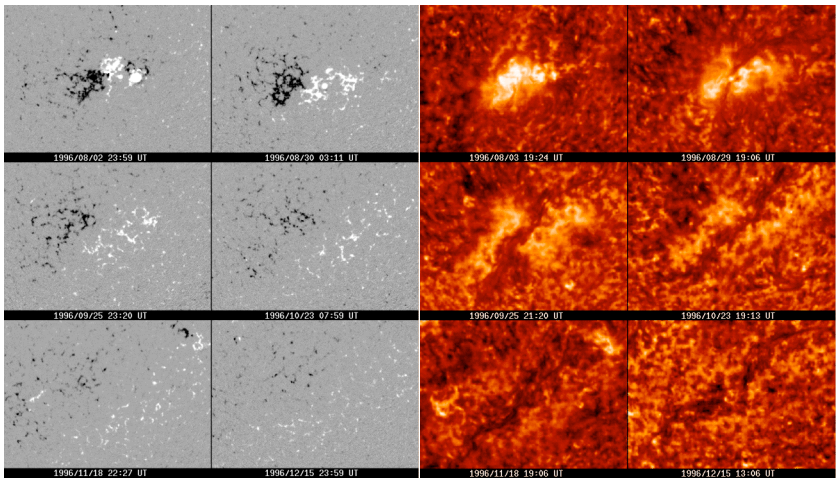
Emergence of NOAA 7978



SOHO/MDI magnetograms

UCL

Evolution of NOAA 7978

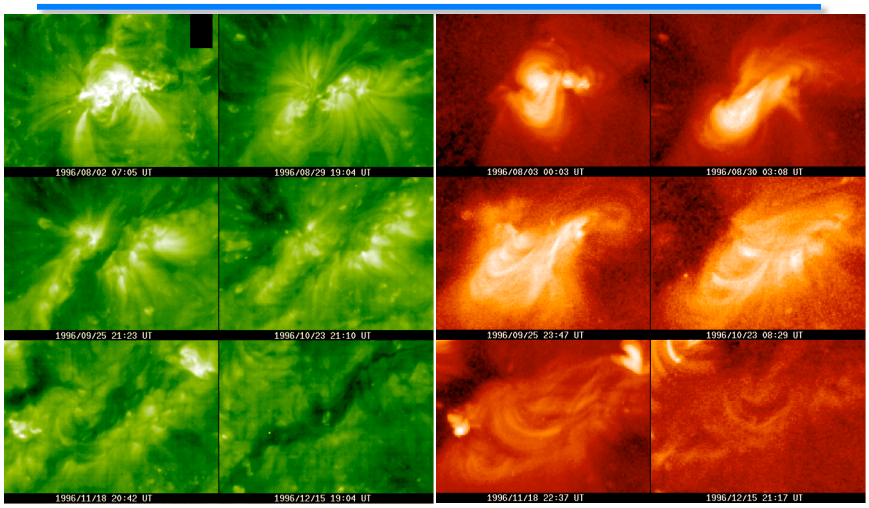


SOHO/MDI
magnetogram

SOHO/EIT
HE II, 304 Å

UCL

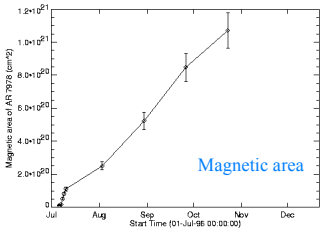
Evolution of NOAA 7978 -- II.

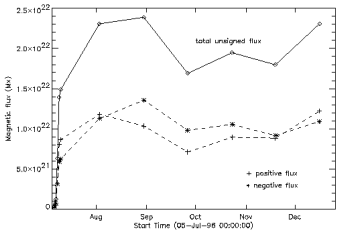


SOHO/EIT
FeXII – 171 Å

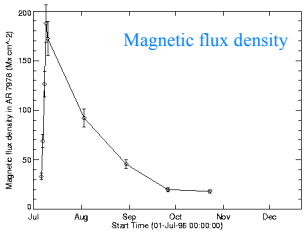
YOHKOH/SXT

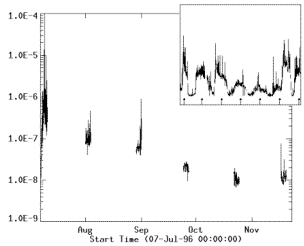
Evolution of the magnetic field





Total magnetic flux

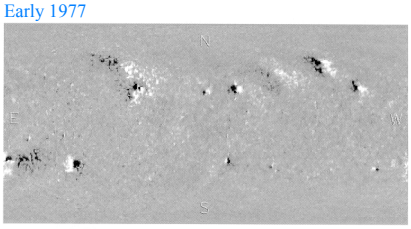




GOES Soft X-ray flux

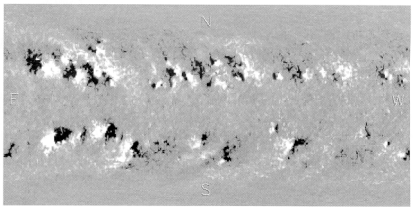
Decay through flux dispersion

Early 1977



Magnetic flux is getting dispersed through a random-walk process (Leighton, 1964)

Lawrence & Schrijver (1993): random walk of particles along random lattices (supergranular network boundaries).



Diffusion or mixing coefficient

D=110-300 km^2/s

found by many studies (quiet sun, active-network, active region)

D=600±200 km^2/s

found from the analysis of large-scale patterns and needed for the dynamo mechanism to work.

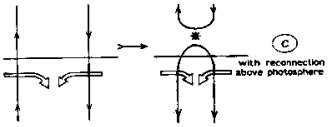
Mid-1978

Removal of magnetic flux

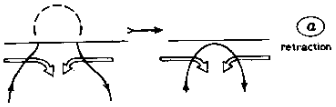
• **Ohmic dissipation** $$\frac{\partial \vec{B}}{\partial t} = \nabla \times (\vec{v} \times \vec{B}) - \eta \nabla \times (\nabla \times \vec{B})$$

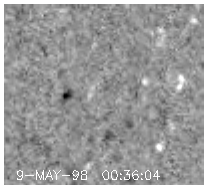
The diffusion timescale is found by equating the LHS of the induction equation with the 2nd term on the RHS: $B/t \approx \eta B/L^2 \Rightarrow t_D \approx L^2/\eta$
That is the time scale of Ohmic decay due to the finite electrical resistance.
For a sunspot with a diameter of $L=10^6$ m, t_D ~ 30,000 year.
Ohmic dissipation becomes effective on the Sun is on sub-granular length scales.

• **magnetic cancellation**
(magnetic reconnection+flux retraction by tension force)

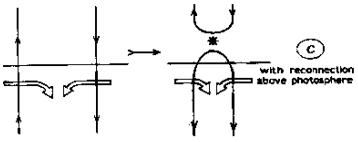


• **flux retraction** (only works for very small elements) large curvature → tension force overcomes buoyancy





Flux cancellation

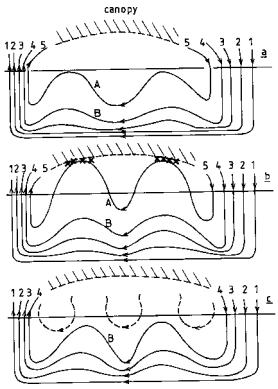


the most common process → associated with X-ray bright points during cancellation.

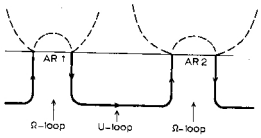
Flux cancellation:
two grains of opposite polarity meet and there is a steady loss of flux in both polarities, until the smallest flux grain has completely disappeared.

In the quiet photosphere the flux-loss rate per pair of magnetic grains is estimated to range from 10^{17} to $4 \cdot 10^{18}$ Mx/hour (Livi et al, 1985).

Flux loss due to U-loop emergence



Spruit et al. (1987)
The sea-serpent process creates a pepper&salt magnetic field and remains unnoticed…



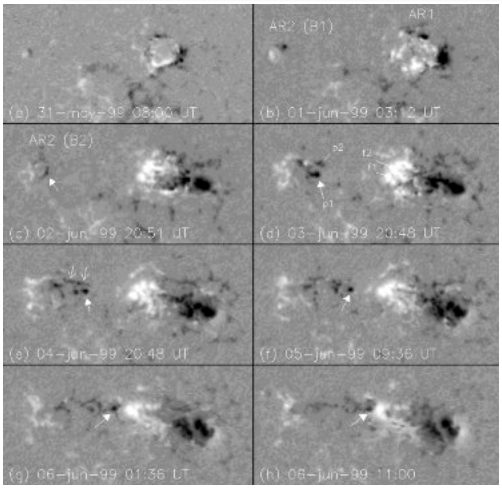
Schrijver & Zwaan (2000)

While **Ω-loops** are indispensable to explain the **emergence** of ARs, **U-loops** are important to understand their **decay** (Zwaan, 1992).

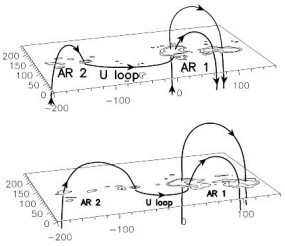
They can be created e.g. **between two adjacent Ω-loops** that have emerged from the **same toroidal flux strand**, or by sub-surface reconnection.
They **have difficulties to emerge** (matter is trapped in their concave-up part, and has to disengage from the plasma (Parker, 1984) → **sea-serpent process** (Spruit et al, 1987).

The existence of U-loops can explain puzzling observations of **in-situ disappearance** of large amount **of flux** from ARs and active nests (instead of submergence).

Emergence of a U-loop connecting two Ω-loops



31 May – 06 June 1999
NOAA 8562 & 8567



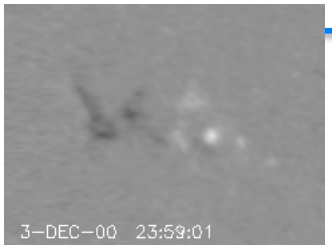
van Driel-Gesztelyi,
Malherbe & Démoulin (2000)

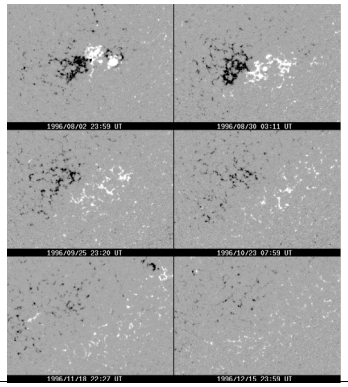
This was a U-loop, because…

- The **proper motion** of the p spots of AR2 indicates a **concave-up flux tube geometry** instead of the usual Ω shape.
- The **elongated shape of the p magnetic field concentrations** of AR2 in the direction of their proper motion indicate a **high inclination** to the vertical.
- The **polarity separation** in AR2 was ≈ 3 times larger than in other ARs with the same flux content (Tian et al. 1999), the balance between buoyancy & tension force was different than in usual Ω-loops.
- The p spot(s) of AR2 and the f spots of AR1 **moved towards one another** both in longitude and latitude – a low-probability event, if accidental.
- When the opposite polarities of the two ARs collided their **cancellation** was **"clean",** it was not accompanied by important flare activity (≤C1.3)

This observation provides the **first direct evidence** for the emergence of a U-loop in the photosphere. (van Driel-Gesztelyi, Malherbe & Démoulin, 2000)

Decay of active regions





MDI magnetograms show the emergence and decay of and active region - note the fast spot decay due to flux bundles streaming out radially from the spot.

$dA/dt \sim -A^{-1/2}$ where A is the area of the umbra.

Evolution of an active region during 6 solar rotations (6 x 28 days):

The flux pealed off the strong flux tubes gets dispersed by random walk process along the supergranular network boundaries.

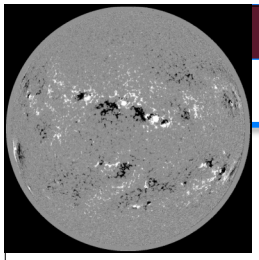
The flux tube of the AR slowly gets disconnected from its toroidal roots:

$$t_{AR}\ (\text{days}) \approx 15\Phi/10^{21}\ \text{Mx}$$

where Φ is the total magnetic flux of the AR ~ 5 months for a major AR (Schrijver & Title, 1999).

Flux-lifetime relationship for AR

Active region	Magnetic flux (Mx)	Lifetime
large AR with sunspots	5×10^{21} - 4×10^{22}	months
small AR with pores - no spots	3×10^{20} - 5×10^{21}	days/weeks
ephemeral AR (X-ray bright points)	2×10^{19} - 3×10^{20}	hours/3 day



Flux budget of the Sun

$\Phi_{SUN} \approx 3 \cdot 10^{23}$ Mx

Magnetic flux scale	Flux emergence rate (Mx/day) solar minimum - maximum	Replacement time solar minimum - maximum
Active region	7×10^{20} - 6×10^{21}	10 months - 4 months
Ephemeral regions	9×10^{22} - 1×10^{23}	2 days - 14 hours
Inter-network flux	10^{24}	5 minutes

- Much more small-scale flux emerges on the Sun than large-scale (AR) flux - they dominate the flux budget.
- However, small-scale flux removed from the surface and replaced quickly, thus the long-surviving large-scale flux defines magnetic face of the Sun.