Stratospheric dynamics, ozone and climate

Theodore G. Shepherd
Department of Physics, University of Toronto

Outline

- The wave-driven meridional circulation
- Stratospheric planetary waves

  → Charney-Drazin theorem
- Polar vortex dynamics

  → Stratospheric sudden warmings
- Brewer-Dobson circulation

  → Seasonal cycle in extratropical ozone
- Chemistry-climate coupling
- Long-term changes in ozone
- Climate change in the stratosphere

The wave-driven diabatic circulation

- Radiation tends to relax the atmosphere towards temperature T_{rad} and zonal wind u_{rad}, in thermal-wind balance
  - → In troposphere, $T_{\text{rad}}/u_{\text{rad}}$ is convectively and baroclinically unstable; leads to vigorous motion
  - → In middle atmosphere, $T_{\text{rad}}/u_{\text{rad}}$ is dynamically stable, so represents a possible quasi-steady state

- Being isolated from the Earth's surface, the middle atmosphere has little thermal inertia
  - → Radiative relaxation time is at most a few weeks
  - → Hence the flow tends to be far more zonally symmetric than the tropospheric flow, making the zonal average more physically meaningful
  - → One expects the lowest temperatures over the winter pole and the highest temperatures over the summer pole
  - → Thus eastward flow in the winter hemisphere, westward flow in the summer hemisphere

- Zonal motion is "free", but rotation and stratification restrict meridional and vertical motion, respectively
  - → Persistent meridional motion requires a torque
  - → Persistent vertical motion requires diabatic heating

- In fact, a persistent meridional circulation is observed

  → Why is it there?

- In the steady limit, the angular momentum balance is

  $f\bar{v}^* \approx -\nabla\cdot\boldsymbol{F}$ (EP flux convergence; a.k.a. "wave drag")

  → Eastward (positive) torque drives equatorward flow, while westward (negative) torque drives poleward flow

  → Angular momentum is *conserved* (even with viscosity)

  → $\nabla\cdot\boldsymbol{F}$ associated with waves; *not* a property of the medium

- In the steady limit, the thermodynamic balance is

$$\frac{N^2T_0}{g}\,\bar{w}^* \approx Q \approx -r(T - T_{\text{rad}})$$

  → Downwelling implies $T > T_{\text{rad}}$, upwelling implies $T < T_{\text{rad}}$

  → Atmosphere is energetically open

- $\bar{v}^*$ and $\bar{w}^*$ constrained by mass continuity, which couples $\nabla\cdot\boldsymbol{F}$ and $T - T_{\text{rad}}$; but which causes which?

- Without wave drag, there would be no radiative heating/ cooling (apart from transient effects) and $T \approx T_{\text{rad}}$

  → Wave drag is a "gyroscopic pump"

  → Radiation doesn't drive the meridional circulation, it *accommodates* it (it's not that "hot air rises")

- Since middle atmosphere is dynamically stable, $\nabla \cdot \boldsymbol{F}$ is associated with waves propagating up from troposphere

  → Unlike troposphere, middle atmosphere is a “refrigerator”

- In the steady limit, can consider $T = \underbrace{T_{\text{rad}}}_{\text{radiation}} + \underbrace{(T - T_{\text{rad}})}_{\text{wave drag}}$

  → In global mean, $\langle T \rangle \approx \langle T_{\text{rad}} \rangle$ at every pressure level

Stratospheric planetary waves

- Consider (for simplicity) the 2-D (barotropic) β-plane eqns

$$\frac{\partial \nabla^2 \psi}{\partial t} + \boldsymbol{v} \cdot \boldsymbol{\nabla}(\nabla^2 \psi + \beta y) = 0 \quad \text{where} \quad \boldsymbol{v} = \left(-\frac{\partial \psi}{\partial y}, \frac{\partial \psi}{\partial x}\right),$$

linearized about a slowly-varying zonal flow $\bar{u}$:

$$\left(\frac{\partial}{\partial t} + \bar{u}\,\frac{\partial}{\partial x}\right)\nabla^2 \psi' + \beta \frac{\partial \psi'}{\partial x} = 0 \quad (\nabla^2 \psi + \beta y = \text{absolute vorticity})$$

→ One obtains the dispersion relation $c = \bar{u} - \dfrac{\beta}{k^2 + \ell^2}$

→ These are *westward* propagating waves (relative to $\bar{u}$)

→ $\beta \sim 10^{-11} \mathrm{m}^{-1} s^{-1}, \quad k \sim \ell \sim 10^{-6} \mathrm{m}^{-1} \quad (\lambda \sim 6000\,\mathrm{km})$

$$\Rightarrow \quad \frac{\beta}{k^2 + \ell^2} \sim 10\,\mathrm{m\,s}^{-1}, \text{ comparable to } \bar{u} \sim 10\,\mathrm{m\,s}^{-1}$$

→ Hence large-scale stationary forcing in an eastward flow efficiently generates planetary-scale Rossby waves

- Now, $\operatorname{sgn}(\overline{u'v'}) = -\operatorname{sgn}(k\ell)$ and $c_{g(y)} = \dfrac{2\beta k\ell}{(k^2+\ell^2)^2}$

  $\Rightarrow$ meridional momentum flux opposite to energy flux

  $\Rightarrow$ Rossby waves carry negative momentum

- There is another way of seeing this

  → Consider the 2-D zonal-mean momentum equation

  → Pressure-gradient and Coriolis terms drop out, leaving

$$\frac{\partial \bar{u}}{\partial t} = -\frac{\partial}{\partial y}(\overline{u'v'}) = \overline{\psi'_{yy}\psi'_x} + \overline{\psi'_y\psi'_{xy}} = \overline{\nabla^2\psi'\psi'_x} - \overline{\psi'_{xx}\psi'_x}$$

$$= -\frac{1}{\beta}\overline{\nabla^2\psi'\Big(\frac{\partial}{\partial t} + \bar{u}\frac{\partial}{\partial x}\Big)\nabla^2\psi'} = \frac{\partial}{\partial t}\Big(-\underbrace{\frac{1}{2\beta}\overline{(\nabla^2\psi')^2}}_{\equiv A}\Big)$$

  → So as Rossby waves enter a region, they decelerate the flow; as they leave it, they accelerate it

  → This effect is illustrated in a laboratory experiment by Whitehead (1975 *Tellus*)

- All this goes through for 3-D (baroclinic) flow, where A is the *Eliassen-Palm wave activity* (a.k.a. negative of the *pseudomomentum*), and for stationary waves

$$\bar{u} = \frac{\beta}{k^2 + \ell^2 + (f^2m^2/N^2)}$$

- → No propagation is possible in the summer stratosphere where $\bar{u} < 0$
- → For $\bar{u}(y,z)$ (slowly varying), ℓ and m will evolve as a wave packet propagates, while k is fixed, but clearly

$$\bar{u} < U_c(k) \equiv \frac{\beta}{k^2}$$

- → Hence if u becomes more positive with altitude, as in the winter stratosphere, the higher-k waves become evanescent
- → Zonal wave 2: $k \sim \dfrac{2\pi}{12\,000\,\mathrm{km}} \quad \Rightarrow \quad U_c \sim 45\,\mathrm{m\,s^{-1}}$
- → Zonal wave 3: $k \sim \dfrac{2\pi}{8\,000\,\mathrm{km}} \quad \Rightarrow \quad U_c \sim 20\,\mathrm{m\,s^{-1}}$
- → Only planetary waves 1 and 2 reach the stratosphere (Charney-Drazin theorem)

- We are now in a position to understand the stratospheric meridional circulation
  - → Persistent planetary-wave forcing gives a negative angular momentum forcing in the winter hemisphere
  - → Negative angular momentum forcing must translate into a poleward mass flux

→ Circulation must return equatorward somewhere, but needs an opposite-signed momentum forcing to do so

→ To the extent that wave momentum transfer is given, this can only happen in the planetary boundary layer ("downward control")

→ This drives upwelling in the tropics, and downwelling in the extratropics, below the forcing level

- This understanding of the meridional circulation explains why the circulation is dominantly in the winter hemisphere

- It also explains the NH-SH asymmetry

  → The NH generates stronger planetary Rossby waves than does the SH

  → The lack of large continental land masses in the SH leads to much lower amplitude planetary waves

  → This leads to a stronger poleward mass flux in the NH than in the SH

  → It also implies more tropical upwelling, and lower tropical lower stratospheric temperatures, in NH winter as compared with SH winter

  → Accounts for the "tropical tape recorder" in water vapour

Polar vortex dynamics

- Polar downwelling from planetary-wave drag warms and weakens the winter polar vortices from radiative equilibrium
  - → More so in the NH than in the SH
  - → Stronger downwelling implies higher polar T, so a weaker meridional gradient of T, so a weaker eastward circumpolar flow
- The stronger Arctic planetary-wave forcing leads to greater variability in the Arctic; there is a nonlinear feedback
  - → Virtually no variability in the quiescent summertime
- Sometimes the forcing is so strong that polar temperatures rise by several tens of degrees K in a few days, and the vortex becomes westward
  - → Called a "stratospheric sudden warming"
  - → Cannot be explained by radiation or horizontal advection
  - → Rare in the SH; first ever observed was in 2002
- Sudden warmings can also be understood as the balanced response to negative angular momentum deposition by planetary waves

→ Need to consider the transient rather than the steady-state response

- Where planetary waves are dissipated they exert a westward torque

  → The torque cannot go entirely into a westward acceleration, because this would violate thermal-wind balance

  → The response goes partly into zonal-wind deceleration, and partly into an instantaneous meridional circulation

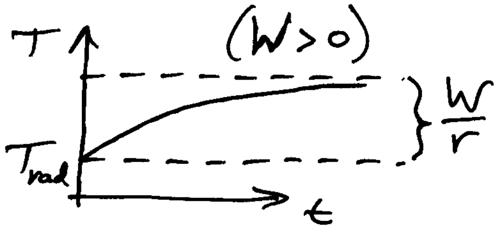
  → Alternatively, the torque goes partly into the relative and partly into the planetary angular momentum

  → The meridional circulation induces a temperature tendency through adiabatic heating/cooling which is in thermal-wind balance with the zonal-wind tendency

  → For fixed forcing this leads to a steady response:

$$\frac{\partial T}{\partial t} + r(T - T_{\mathrm{rad}}) = -\frac{N^2 T_0}{g}\,\bar{w}^* \equiv W = \mathrm{const.}$$

$$\Longrightarrow \quad T = T_{\mathrm{rad}} + \frac{W}{r}(1 - \mathrm{e}^{-rt})$$



- Normally, stationary planetary waves propagate towards the equator and dissipate in midlatitudes

→ Sudden warmings occur when the wave dissipation manages to penetrate into the Arctic vortex, and the induced downwelling is focused over the pole

→ When the vortex becomes westward, wave propagation shuts down and the vortex recovers radiatively

- There is another interpretation of sudden warmings in terms of 2-D vortex dynamics (corresponding to an isentropic layer)

→ We know that absorption of Rossby waves must lead to a reduction in angular momentum

→ In planar geometry, angular momentum is given by

$$\iint \hat{\boldsymbol{z}} \cdot (\boldsymbol{r} \times \boldsymbol{v})\,\mathrm{d}x\mathrm{d}y = -\frac{1}{2}\iint r^2\omega\,\mathrm{d}x\mathrm{d}y \quad (\omega = \text{vorticity})$$

→ In spherical geometry, it is given by

$$\iint \sin\phi\, q\, \mathrm{d}(\sin\phi)\mathrm{d}\lambda \quad (q = \text{potential vorticity});$$

note that $\sin\phi \approx 1 - \frac{1}{2}(\phi')^2$ near the pole, $\phi' \equiv \frac{\pi}{2} - \phi$

→ Since either ω or q are materially conserved, a reduction in angular momentum with the same ω or q requires a *deformation* of the vortex

→ And indeed sudden warmings are associated with a break-up of the polar vortex

Brewer-Dobson circulation

- Contrast in mixing timescales between troposphere and stratosphere ⇒ tropopause is a distinct chemical boundary for long-lived species
  - → Analogous to top of the planetary boundary layer
- Chemical measurements have always played a crucial role in understanding the stratospheric circulation
  - → Within the stratosphere, the mean diabatic circulation also transports chemical species
  - → Poleward sense of circulation originally inferred from water vapour (Brewer 1949) and ozone (Dobson 1956)
  - → Clearly visible in long-lived species, e.g. CH_4 and N_2O
- Implies that the tropics are the entry point for the stratosphere, and the extratropics the departure point
  - → Diabatic (mass) circulation represents the “advective” part of the flux of species
  - → Has overturning time scale of several years
  - → Acts to steepen meridional gradients of long-lived species

- There is also a contribution due to mixing (with no net mass flux)
  - → For planetary waves, it is closely coupled to transport since $\nabla \cdot F = \overline{v'q'}$ (q is potential vorticity)
  - → Acts to flatten meridional gradients of long-lived species
  - → Observed distributions represent a balance between these two effects

- Because of mixing, there is a distribution of transit times to any particular region of the stratosphere
  - → Mean of the distribution is the "mean age"; oldest in polar upper stratosphere

- Horizontal mixing is strongest in the "surf zone" of the winter hemisphere, associated with breaking planetary waves
  - → Surf zone q stirring drives the Brewer-Dobson circulation
  - → Deformation of q is associated with absorption of the negative angular momentum of the planetary waves, as with the vortex break-up in sudden warmings
  - → Horizontal inhomogeneity of mixing leads to "mixing barriers": polar vortex edge and tropical pipe

- Seasonality of stratospheric planetary-wave drag implies a seasonality in the Brewer-Dobson circulation

  → Leads to spring buildup of extratropical column O_3 (more so in NH than in SH)

  → Ozone decays photochemically in the quiescent summer

  → Summertime ozone more similar between NH and SH

- Interannual variability of planetary-wave drag implies interannual variability in the wintertime ozone buildup

  → Variations in tropical and extratropical O_3 anti-correlated

  → More O_3 buildup in spring implies more chemical O_3 loss in summer — you can only destroy what you have

  → There is a remarkable seasonal persistence in the interannual ozone anomalies

  → Midlatitude and polar variations are highly correlated

Chemistry-climate coupling

- Relationship between diabatic circulation and transport, via wave drag, is crucial to understanding cause and effect

  → Important for attribution of observed record

  → Important for diagnosing causes of model errors

An example: If stratospheric wave drag decreased then:

→ Diabatic and Brewer-Dobson circulations would weaken

→ Tropical tropopause would warm, hence more stratospheric H_2O

→ Air would be older, hence also more stratospheric H_2O (more CH_4 oxidation) and less CH_4

→ More HO_x, hence more chemical ozone loss (except in middle stratosphere)

→ Less diabatic descent in midlatitudes and poles, hence colder temperatures

→ Less ozone transport to midlatitudes and poles, hence still colder temperatures

→ Stronger polar vortex, hence more of a barrier to ozone transport, and less polar ozone

→ Increased PSC chemistry, hence more chemical ozone loss, and still less polar ozone

- On long timescales, changes in atmospheric composition can change both T_{rad} and wave drag

- But since radiative and chemical effects are dissipative (stable), unforced natural variability arises solely from dynamics, through variability in wave drag

  → Variability in tropospheric forcing of planetary waves

  → Variability internal to the stratosphere (e.g. QBO)

- There are feedbacks, which tend to be positive:

  → More transport ⇒ more O_3 ⇒ still more heating

  → Higher T ⇒ less heterogeneous chemistry ⇒ still more O_3

  → Weaker vortex ⇒ more transport ⇒ still more O_3

- So one can think of chemistry and radiation as amplifiers of the dynamical variability, but the variability originates in the dynamics (apart from volcanic eruptions)

- Link between diabatic and Brewer-Dobson circulations leads to positive O_3–T correlation in the lower stratosphere — *which is, in itself, purely coincidental*

  → Extent of coupling, and causality of statistical relationships, need to be assessed on a case-by-case basis

Long-term changes in ozone

- Chlorine loading over last few decades has led to Antarctic ozone hole
  - → Antarctic is always cold enough for PSC formation
  - → Indeed even before the chlorine build-up and chemical ozone loss, the Antarctic exhibited a transport-induced, purely dynamical, "ozone hole"
- No ozone hole in Arctic because vortex is too warm
  - → When the Arctic has a cold winter, then chemical ozone loss occurs
  - → However low T from less downwelling ⇒ less O_3 transport
  - → Low O_3 in recent cold Arctic winters comes half from chemical loss, half from reduced transport (WMO 2003)
- In Arctic, ozone levels controlled by meteorology more than by halogen loading (but effects are coupled)
  - → One has to look at effects of halogen loading on ozone variability
  - → Observed record is just one realization from an ensemble of possibilities

- Chemistry climate models reproduce development of Antarctic ozone hole, but don't simulate the Arctic record particularly well (whereas CTMs do, since driven by obs)

  → But does this mean they are wrong?

- There have also been ozone decreases in midlatitudes

  → Can be simulated by 2D models, but there are issues

We can only understand the ozone record (outside of Antarctica) in the context of the dynamical record

- There have been long-term changes in various stratospheric climate indicators over past decades

  → Arctic wintertime vortex has gotten colder and stronger, and more persistent

  → Stratopheric air has become "older"

  → Stratospheric PWD has weakened

- These results are all qualitatively consistent with each other

- However they are quite sensitive to the months and time period considered

  → Arctic exhibited warming (and more PWD) in early winter, and cooling (and less PWD) in late winter

  - → Note that this cooling cannot be explained by ozone loss

- A long-term decrease in stratospheric PWD implies a long-term decrease in the winter-to-summer O_3
  - → Over 1979-2000, estimate is about 20-30% of midlatitude total O_3 changes in January-March (WMO 2003)
  - → Is there a chemical amplification, as in the Arctic?

Climate change in the stratosphere

- WMGHGs will increase as halogens decrease
  - → Climate change and ozone recovery problem are coupled
  - → Furthermore ozone is a GHG, which interacts strongly with temperature especially in the upper stratosphere

- Increasing CO_2 will cool the stratosphere radiatively, but mainly the upper stratosphere
  - → Increased IR emission exceeds increased IR absorption
  - → Decreases O_3 loss rates, hence enhances O_3 recovery

- In the lower stratosphere, cooling could potentially *decrease* O_3, but direct radiative effect is weak

→ The most significant effects would arise from dynamical feedbacks (meaning planetary wave drag)

- GCM predictions of greenhouse-gas-induced changes in NH stratospheric wave drag diverge widely
  - → Older results tended to suggest a reduction in NH stratospheric wave drag (Shindell *et al.* 1998 *Nature*)
  - → Was argued by some to explain observed Arctic trends in 1990s (and to imply a future Arctic O_3 hole) — but not after 1998!
  - → More recent studies tend to suggest an *increase* in stratospheric wave drag

- We probably cannot say anything with any confidence
  - → This aspect of GCM behaviour is highly sensitive
  - → In most cases, results are not statistically significant
  - → Certainly GCMs do not explain the past NH behaviour
  - → Long-term records show no particular trend

- Natural variability may well be the dominant factor in the evolution of Arctic ozone over the next few decades
  - → However the Antarctic ozone hole should recover by $\sim$ 2050 (will take a decade or so to see turnover)