**Physics of Atmospheres and Oceans: Class Question Sheets**

REMOTE SOUNDING

**RS.1**

*The figure shows emission spectra from Mars obtained by Hanel* et al. *with the IRIS instrument on Mariner 9.*

*(i) What do the fine features between 200 and 400 cm$^{-1}$, especially in the south polar spectra, indicate?*

The presence of Water Vapour [this is a water vapour rotation band].

[1]

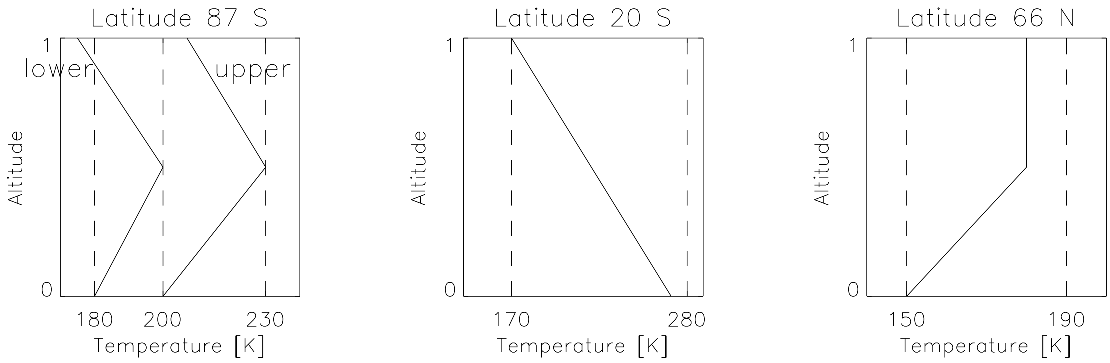
*(ii) What is responsible for the large feature centred on 667 cm$^{-1}$?*

Carbon Dioxide [$CO_2$ $\nu_2$ (bending mode) vibration-rotation band]

[1]

*(iii) What can be said about the surface and atmospheric temperatures in each case? Sketch temperature profiles that would give rise to spectra like these.*

Assume the atmosphere is opaque at high altitudes at the centre of the $CO_2$ band, opaque at low altitudes at the sides of the band, and transparent outside the band. Temperatures at each level are then obtained by comparing the emission with the Planck Function.




[5]

*(iv) What can be learnt from a comparison of the spectra for terrestrial mid-latitudes (e.g. Houghton 2nd edition fig. 12.6(b), 1st edition 12.3(b)) and for Mars at 20° S?*

Mars has little or no $O_3$, $CH_4$ or $N_2O$, and a small amount of $H_2O$.

[3]

[Also, further information which may be noted:
(a) There are occasional broad features around 1100cm$^{-1}$ due to $SiO_2$, i.e. dust storms
(b) Mars is colder than the Earth (both surface and atmosphere)
(c) 87°S surface temperature is possibly non-uniform (surface spectra do not lie along the same temperature curve either side of the $CO_2$ feature)
(d) Atmosphere cools between acquisition of upper and lower spectra at 87°S.
(e) More water vapour in the south polar summer, suggesting that the source is the melting polar cap.]

Total: [10]

**RS.2**

*Measurements of outgoing radiation are made in three spectral intervals by a nadir viewing radiometer in Earth orbit, with the following results:*

*Interval 1: brightness temperature = 310 K (atmospheric window at 900 $cm^{-1}$)*

*Interval 2: brightness temperature = 220 K ($CO_2$ band at 667 $cm^{-1}$)*

*Interval 3: brightness temperature = 280 K ($O_3$ band at 1040 $cm^{-1}$)*

*Brightness temperature is the temperature that would be inferred from a radiometric measurement by assuming that the source is black.*

*Assume that the surface is black, that the $CO_2$ absorption is so strong in interval 2 that the stratosphere is black in that interval, and that all the ozone is in the stratosphere, calculate the transmission of the ozone layer at 1040 $cm^{-1}$. What could we learn about the ozone layer from this?*

*[The Planck spectral radiance at 1040 $cm^{-1}$ is 14.9, 64.3 and 108.2 mW $m^{-2} ster^{-1}/cm^{-1}$ at 220, 280 and 310 K respectively.]*

Since the stratosphere is opaque in the $CO_2$ band, the upper stratosphere temperature can be taken as the $CO_2$ brightness temperature: 220 K. Similarly, the surface temperature can be taken as the window brightness temperature: 310 K.

In general, the upward radiance $R$ measured at the top of an atmosphere above a black surface is given by:

$$R = B(T_s)\tau_s + \int_0^\infty B(T_z)\frac{d\tau}{dz}\,dz$$

where $T_s$ is the surface temperature, $T_z$ is the temperature at some level $z$ in the atmosphere, and $\tau$ is the transmittance of the atmosphere ($\tau = 1$ at the top of the atmosphere, $\tau = \tau_s$ at the surface). For the $O_3$ band we are given $R = B(280)$. If it is assumed that most of the $O_3$ is confined to the upper stratosphere, the $B(T_z)$ term can be replaced by $B(220)$ and removed from the integration, leaving:

$$\begin{aligned} B(280) &= B(310)\tau_s + B(220)\int_0^\infty \frac{d\tau}{dz}\,dz \\ &= B(310)\tau_s + B(220)(1-\tau_s) \end{aligned}$$

Rearranging this gives:

$$\tau_s = \frac{B(280)-B(220)}{B(310)-B(220)} = \frac{64.3-14.9}{108.2-14.9} = 0.53$$

[3]

The (monochromatic) transmittance is related to the ozone density distribution $\rho(z)$ by:

$$\tau_s = \exp\left(-\int_0^\infty k\rho\,dz\right)$$

where $k$ is the absorption coefficient (e.g. measured in $m^2$/mole of $O_3$). If it is assumed that $k$ is known (and independent of temperature or pressure) then the measurement of transmittance can be used to derive the total column amount of ozone $u = \int \rho\,dz$ (e.g. mole/$m^2$).

[2]

Total: [5]

**RS.3**

*The 11 $\mu$m region is often referred to as a "window" in the terrestrial atmosphere, but there is still some absorption mostly due to water vapour. Assuming that the mass mixing ratio $x$ at pressure level $p$ is $x_0(p/p_0)^3$ and is $\ll 1$, where $x_0$ and $p_0$ refer to the bottom boundary, and that the absorption coefficient $k$ due to water is $k = \alpha p_w$ where $p_w$ is the vapour pressure and $\alpha$ is a constant, show that the transmittance of a vertical path from level $p$ to a satellite is $\exp(-\beta(p/p_0)^8)$ where*

$$\beta = \frac{\alpha}{g}\frac{M}{M_w}\frac{x_0^2 p_0^2}{8}$$

*and $M_w$ and $M$ are the molar masses of water vapour and air respectively.*

Transmittance from level $z$ is given by:

$$\tau = \exp\left(-\int_z^\infty k\rho_w\,dz\right)$$

Expressing the various components as functions of pressure:

$$k = \alpha p_w = \alpha x p\left(\frac{M}{M_w}\right) = \alpha x_0 p\left(\frac{p}{p_0}\right)^3\left(\frac{M}{M_w}\right)$$

[$M/M_w$ arises because $x$ is a *mass* mixing ratio, not a *volume* mixing ratio].

$$\rho_w\,dz = x\rho\,dz = -x\frac{dp}{g} = -\frac{x_0}{g}\left(\frac{p}{p_0}\right)^3 dp$$

[hydrostatic equation $dp = -g\rho\,dz$ requires air density $\rho$ — *not* water density $\rho_w$ — hence factor $x$]

Combining these terms gives the required answer [limits of integration are switched, corresponding to change of sign from $z \to p$]:

$$\tau = \exp\left(-\int_0^p \frac{\alpha}{g}\frac{M}{M_w}x_0^2\left(\frac{p}{p_0}\right)^6 p\,dp\right) = \exp\left(-\beta\left(\frac{p}{p_0}\right)^8\right)$$

[5]

*Given that $\alpha = 7\times10^{-6}$ $m^2kg^{-1}Pa^{-1}$ and that the temperature $T_0$ and relative humidity $\Lambda_0$ at the bottom boundary are 290 K and 70% respectively, show that $\beta = 0.102$ and calculate the transmittance from the surface to the satellite. [The SVP of water vapour at 290 K is 1936.7 Pa.]*

At surface, the partial pressure due to water vapour, $e_0 = \Lambda_0 e_{\rm SVP}$ [$e_0 = 0.7\times1936.7 = 1355.7$ Pa]. The surface value of mass mixing ratio $x_0$ is given by $x_0 = (e_0 M_w)/(p_0 M)$ [$x_0 = 8.43\times10^{-3}$]. Replacing $x_0$ in the expression for $\beta$ gives:

$$\beta = \frac{\alpha}{g}\frac{M_w}{M}\frac{(\Lambda_0 e_{\rm SVP})^2}{8} = \left(\frac{7\times10^{-6}}{9.81}\right)\left(\frac{18}{28.96}\right)\frac{(0.7\times1936.7)^2}{8} = 0.102$$

$$\tau = \exp\left(-\beta\left(\frac{p}{p_0}\right)^8\right) = \exp(-\beta) = 0.903$$

[3]

*Write down an expression for the net deficit in the spectral radiance received from the surface by a nadir viewing radiometer due to attenuation and emission by water vapour. Evaluate this deficit as a fraction of $B(T_0)$ for the case of a (black) sea surface at temperature $T_0$ for the conditions given above assuming that $T = T_0(p/p_0)^\delta$, $B = \eta(T/300)^\mu$. Take $p_0 = 10^5$ Pa, $\delta = 0.29$, $\eta = 0.12$ $Wm^{-2}ster^{-1}/cm^{-1}$ and $\mu = 4.5$.*

$$\left[\int_0^1 y^{8.305} \exp(-0.102y^8)\, dy = 0.102\right]$$

The (positive) radiance 'Deficit' $\Delta B$ is defined as the difference between the surface-emitted radiance $B(T_0)$ and the observed radiance $L$.

$$\Delta B = B(T_0) - L = B(T_0) - \left(B(T_0)\tau_0 + \int_{p_0}^{0} B(T,p)\frac{d\tau}{dp}\, dp\right) = B(T_0)(1-\tau_0) - \int_{p_0}^{0} B(T,p)\frac{d\tau}{dp}\, dp$$

[2]

Using given expressions for $T(p)$, $B(T)$, and substituting $y = p/p_0$ so that $T = T_0 y^\delta$, $B(T) = B(T_0)y^{\mu\delta}$, $\tau = \exp(-\beta y^8)$, this becomes:

$$\begin{aligned}
\Delta B &= B(T_0)(1-\tau_0) - \int_1^0 B(T_0)y^{\mu\delta}\left(-8\beta y^7 \exp(-\beta y^8)\right)\, dy \\
\frac{\Delta B}{B(T_0)} &= 1 - \tau_0 + 8\beta \int_1^0 y^{7+\mu\delta} \exp(-\beta y^8)\, dy \\
&= 1 - 0.903 + 8 \times 0.102 \int_1^0 y^{8.305} \exp(-0.102y^8)\, dy \\
&= 0.097 - 0.083 = 0.014
\end{aligned}$$

[5]

*What would be the error in inferring the temperature of the sea surface from radiance measurements if the effect of the water vapour were not taken into account? Is this error significant?*

Differentiating $B = \eta(T/300)^\mu$ [although easier to take log first, then differentiate: $\ln B = \mu \ln T + \text{constant}$]

$$\frac{dB}{dT} = \frac{\mu\eta}{300}\left(\frac{T}{300}\right)^{\mu-1} = \mu\frac{B}{T}$$

Rearrange this to give a temperature deficit $\Delta T$ resulting from a radiance deficit $\Delta B$:

$$\Delta T = \frac{T}{\mu}\frac{\Delta B}{B} = \frac{290}{4.5} \times 0.014 = 0.90 \text{ K}$$

[NB: value of $\eta$ is not required].

[3]

[Alternatively, going the long way round: for $T$=290 K, $B = \eta(T/300)^\mu = 0.103$ W m$^{-2}$ ster$^{-1}$/cm$^{-1}$]:

$$\Delta T = \frac{300}{\mu\eta}\left(\frac{300}{T}\right)^{\mu-1}\Delta B = \frac{300}{4.5 \times 0.12}\left(\frac{300}{290}\right)^{3.5}(0.014 \times 0.103) = 555.6 \times 1.126 \times 1.44 \times 10^{-3} = 0.90 \text{ K}$$

Typical seasonal variability of Sea Surface Temperature at any location is only a few K, so, to be useful, SST measurements have to have an absolute accuracy better than $\pm 0.5$ K. Therefore such a temperature deficit is significant.

[2]

[Determining SST from satellites requires the use of different angles and/or spectral channels to correct for this atmospheric absorption. Target accuracies of these SST measurements are of the order of 0.1–0.2 K.]

Total: [20]

**RS.4**

*What is meant by a* weighting function *in atmospheric remote sounding?*

The weighting function is the rate of change of transmittance with height $d\tau/dz$ [p.14 lecture notes], or with respect to some other vertical coordinate.

[Strictly speaking, the weighting function is the sensitivity of a measurement to the retrieved atmospheric parameters, which is only equivalent to $d\tau/dz$ if Planck function profile $B(z)$ is being retrieved].

[2]

*A temperature sounder measures the thermal radiation emitted vertically from an atmosphere for which the transmittance from pressure level p to space is $\tau(p) = \exp(-\beta p^\alpha)$, where $\alpha$ and $\beta$ are constants. Using the height-like variable $\zeta = -\ln(p/p_0)$ as a vertical coordinate, show that the weighting function for Planck function is*

$$K(p) = \alpha\beta p^\alpha \exp(-\beta p^\alpha)$$

Expressing the standard upward radiance equation in $\zeta$-coordinates, and assuming the atmosphere is opaque (so no surface emission term):

$$L = \int_0^\infty B(T,\zeta)\frac{d\tau}{d\zeta}\, d\zeta$$

The weighting function $K$ can be identified as the $d\tau/d\zeta$ term:

$$K = \frac{d\tau}{d\zeta} = \frac{d\tau}{dp} \times \frac{dp}{d\zeta} = -\alpha\beta p^{\alpha-1}\exp\left(-\beta p^\alpha\right) \times (-p) = \alpha\beta p^\alpha \exp\left(-\beta p^\alpha\right)$$

[Note: $\zeta = -\ln(p/p_0) = -\ln p + \ln p_0$, so that $d\zeta/dp = -1/p$]

[2]

*Find $\alpha$ and $\beta$ for an isothermal atmosphere containing an absorber with a small and constant mass mixing ratio $x$ when*

*(i) The absorber is grey with constant absorption coefficient $k$.*

The standard expression for transmittance for an absorber density $\rho_a$ is $\tau = \exp(-\int k\rho_a dz)$. Thus $\beta p^\alpha$ is equivalent to $\int k\rho_a dz$. The absorber density $\rho_a$ is related to the air density $\rho$ by the mass mixing ratio: $x = \rho_a/\rho$. The altitude coordinate $dz$ can be converted to pressure using the hydrostatic equation: $dp = -g\rho dz$. So:

$$\int_z^\infty k\rho_a dz = \int_0^p k\,x\rho\,\frac{dp}{g\rho} = \frac{kx}{g}\int_0^p dp = \frac{kxp}{g} \equiv \beta p^\alpha$$

[Note switch of integration limits $(z,\infty) \rightarrow (0,p)$ which absorbs the negative sign in the hydrostatic equation]. Giving: $\alpha = 1$ and $\beta = kx/g$

[2]

*(ii) The radiometer is sensitive to a single frequency in the far wings of a Lorentz line (ignore temperature dependence of the line parameters).*

The Lorentz line shape is given by:

$$k = \frac{S\gamma_L}{\pi\left[(\nu-\nu_0)^2 + \gamma_L^2\right]} \simeq \frac{S\gamma_L}{\pi(\nu-\nu_0)^2}$$

Where it has been assumed that $\nu - \nu_0 \gg \gamma_L$ in the far wings. The Lorentz halfwidth $\gamma_L$ varies with pressure according to $\gamma_L = \gamma_0(p/p_0)$, so substituting this expression for $k$ (now a function of $p$) in the same integration as before:

$$\frac{xS\gamma_0}{g\pi p_0(\nu-\nu_0)^2}\int_0^p p\,dp = \frac{xS\gamma_0}{2g\pi p_0(\nu-\nu_0)^2}\,p^2 \equiv \beta p^\alpha$$

Giving: $\alpha = 2$ and $\beta = xS\gamma_0/2g\pi p_0(\nu-\nu_0)^2$.

[2]

*(iii) The radiometer is sensitive to a spectral interval in which the transmittance from level p to the satellite is given by the strong limit of the Goody random model,*

$$\tau = \exp[-\frac{2}{\delta}(S\int_p^\infty \gamma_L(z)x\rho(z)dz)^{\frac{1}{2}}]$$

*where $S$ is mean line strength, $\rho$ is air density, $\gamma_L$ is the Lorentz halfwidth and $\delta$ is the mean line spacing.*

Following the same procedure:

$$\frac{2}{\delta}\left(S\int_0^p \frac{\gamma_0 p}{p_0}x\frac{dp}{g}\right)^{\frac{1}{2}} = \frac{2}{\delta}\left(\frac{xS\gamma_0}{gp_0}\int_0^p p\,dp\right)^{\frac{1}{2}} = \frac{2}{\delta}\left(\frac{xS\gamma_0 p^2}{2gp_0}\right)^{\frac{1}{2}} = \frac{1}{\delta}\left(\frac{2xS\gamma_0}{gp_0}\right)^{\frac{1}{2}} p \equiv \beta p^\alpha$$

Giving $\alpha = 1$ and $\beta = (2xS\gamma_0/\delta^2 gp_0)^{1/2}$.

[2]

Total: [10]

**RS.5**

*Find the pressure levels at which the peaks of the weighting functions occur, and find also the widths (in terms of log pressure) at half maximum of the weighting functions (i) and (ii) of the previous question. Sketch the functions for these cases.*

*[The roots of $2x = e^{x-1}$ are approximately 2.68 and 0.23.]*

Differentiating the general form $K = \alpha\beta p^\alpha \exp(-\beta p^\alpha)$:

$$\frac{dK}{dp} = \frac{\alpha}{p}K - \alpha\beta p^{\alpha-1}K = 0 \quad \text{at } p_{\max}$$

Multiplying by $p_{\max}/\alpha K$:

$$\begin{aligned} 1 - \beta p_{\max}^\alpha &= 0 \\ &\Rightarrow p_{\max} = \beta^{-1/\alpha} \\ &\Rightarrow \zeta_{\max} = -\ln(p_{\max}/p_0) = \frac{1}{\alpha}\ln\beta + \ln p_0 \\ &\Rightarrow K_{\max} = \alpha\beta p_{\max}^\alpha \exp(-\beta p_{\max}^\alpha) = \alpha\exp(-1) \end{aligned}$$

[2]

To determine $\zeta$ for which $K$ falls to half maximum $\alpha\exp(-1)/2$:

$$\frac{\alpha}{2}\exp(-1) = \alpha\beta p^\alpha \exp(-\beta p^\alpha)$$

Multiplying by $2\exp(\beta p^\alpha)/\alpha$:

$$\exp(\beta p^\alpha - 1) = 2\beta p^\alpha$$

Identifying $x \equiv \beta p^\alpha$, this has roots 2.68 and 0.23, therefore

$$p = \left(\frac{2.68}{\beta}\right)^{1/\alpha}, \quad \left(\frac{0.23}{\beta}\right)^{1/\alpha} \quad \Rightarrow \quad \zeta_{1/2} = -\frac{1}{\alpha}\ln\left(\frac{2.68}{\beta}\right) + \ln p_0, \quad -\frac{1}{\alpha}\ln\left(\frac{0.23}{\beta}\right) + \ln p_0$$

The width at half-maximum $\Delta\zeta$ is the difference between the two solutions for $\zeta_{1/2}$:

$$\Delta\zeta = \frac{1}{\alpha}\left(\ln\left(\frac{2.68}{\beta}\right) - \ln\left(\frac{0.23}{\beta}\right)\right) = \frac{1}{\alpha}\ln\left(\frac{2.68}{0.23}\right) = \frac{2.455}{\alpha}$$

[2]

Substituting the values of $\alpha$ and $\beta$ derived in the previous question:

(i) ($\alpha = 1$, $\beta = kx/g$):

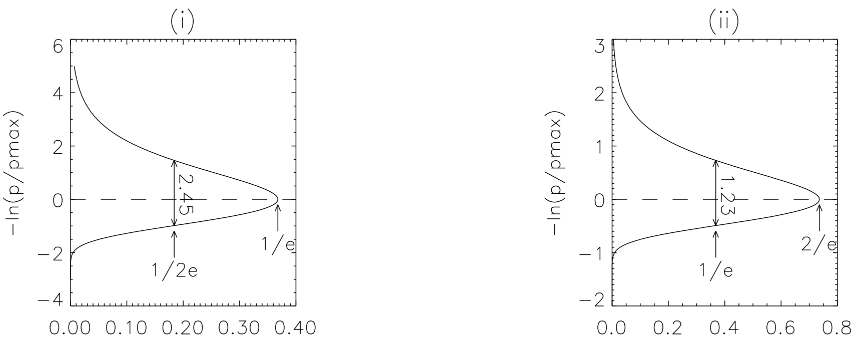
$$p_{\max} = \left(\frac{g}{kx}\right), \qquad \Delta\zeta = 2.455$$

[2]

(ii) ($\alpha = 2$, $\beta = xS\gamma_0/2g\pi p_0(\nu-\nu_0)^2$):

$$p_{\max} = \left(\frac{2g\pi p_0(\nu-\nu_0)^2}{xS\gamma_0}\right)^{\frac{1}{2}}, \qquad \Delta\zeta = 1.228$$

[2]




[2]

Total: [10]

**RS.6**

*A set of remote-sounding measurements* $y_i, i = 1 \ldots m$*, is related to an unknown quantity* $x(z)$*, a function of height* $z$*, by*

$$y_i = \int_0^\infty K_i(z) x(z)\, dz + \epsilon_i$$

*where* $K_i(z)$ *is a set of known functions. What does the term* $\epsilon_i$ *indicate?*

$\epsilon_i$ is the measurement error.

[1]

*Show how this relationship can be put in the form of a matrix equation*

$$\mathbf{y} = \mathbf{K}\mathbf{x} + \boldsymbol{\epsilon}$$

*where* $\mathbf{y}$ *and* $\boldsymbol{\epsilon}$ *are vectors whose elements are* $y_i$ *and* $\epsilon_i$*,* $\mathbf{K}$ *is an* $m \times n$ *matrix, and* $\mathbf{x}$ *is a vector of length* $n$*. (Note that* $\mathbf{K}$ *is not necessarily square.) Indicate any assumption(s) you make, and explain how* $\mathbf{K}$ *and* $\mathbf{x}$ *are related to* $K_i(z)$ *and* $x(z)$*.*

Assumption [p26 of the lecture notes]: $x(z)$ can be represented by a finite set of $n$ functions $W_j(z)$ (e.g. polynomials, Fourier series, interpolation functions), with coefficients $x_j$:

$$x(z) = \sum_{j=1}^{n} x_j W_j(z)$$

Then:

$$y_i = \sum_{j=1}^{n} x_j \int_0^\infty K_i(z) W_j(z)\, dz + \epsilon_i$$

The integrated terms can be represented as elements of a matrix $\mathbf{K}$:

$$K_{ij} = \int_0^\infty K_i(z) W_j(z)\, dz$$

Hence, for measurements $i = 1 \ldots m$,

$$\mathbf{y} = \mathbf{K}\mathbf{x} + \boldsymbol{\epsilon}$$

[2]

*What determined the choice of* $n$*? When is it (i) possible and (ii) desirable to seek a solution* $\hat{\mathbf{x}}$ *which fits the measurement exactly, i.e. one for which* $\mathbf{K}\hat{\mathbf{x}} = \mathbf{y}$*?*

$n$ is usually chosen to be as large as possible to obtain the maximum information from the measurements (e.g. in terms of vertical resolution), but within the requirement that the problem remains well-constrained, which usually means $n \leq m$ [necessary but not sufficient since this also depends on the type of measurements, and the 'number of measurements' $m$ may also include 'virtual' measurements such as additional a-priori constraints].

(i) Possible to seek an exact solution when $n \geq m$

(ii) Desirable only if this produces a solution which is insensitive to measurement noise (which is not usually the case for remote sounding) [exact fitting is more common in, say, ocean modelling where the actual measurements are few and of high quality compared to other constraints such as the initial conditions].

[2]

Total: [5]

**RS.7**

*Obtain solutions to the linear retrieval problem* $\mathbf{y} = \mathbf{K}\mathbf{x} + \boldsymbol{\epsilon}$ *for the following cases: (you may use either component notation or matrix notation in your answer).*

*(a) The exact solution of the underconstrained problem* ($m < n$) *for which* $\mathbf{x}$ *is smallest in the sense that* $\sum_j x_j^2$ *is minimised.*

In vector notation: the cost term is represented by $\sum_j x_j^2 = \mathbf{x}^T\mathbf{x}$. Associating the $m$ constraining equations ($\mathbf{y} = \mathbf{K}\mathbf{x}$ for the exact solution) with a vector of $m$ Lagrange multipliers $\boldsymbol{\lambda}$, and applying the usual minimisation equation in vector form

$$\frac{\partial}{\partial \mathbf{x}}\left(\mathbf{x}^T\mathbf{x} + (\mathbf{y} - \mathbf{K}\mathbf{x})^T\boldsymbol{\lambda}\right) = 0 \quad \text{[NB: } (\mathbf{K}\mathbf{x} - \mathbf{y})^T\boldsymbol{\lambda} \text{ or } \boldsymbol{\lambda}^T(\mathbf{y} - \mathbf{K}\mathbf{x}) \text{ also OK} \rightarrow \text{same solution]}$$

$$2\mathbf{x} - \mathbf{K}^T\boldsymbol{\lambda} = 0$$

$$\mathbf{x} = \frac{1}{2}\mathbf{K}^T\boldsymbol{\lambda}$$

$$\text{substituting into the constraints:} \quad \mathbf{y} = \mathbf{K}\mathbf{x} = \frac{1}{2}\mathbf{K}\mathbf{K}^T\boldsymbol{\lambda}$$

$$\boldsymbol{\lambda} = 2\left(\mathbf{K}\mathbf{K}^T\right)^{-1}\mathbf{y}$$

$$\mathbf{x} = \frac{1}{2}\mathbf{K}^T\boldsymbol{\lambda} = \mathbf{K}^T\left(\mathbf{K}\mathbf{K}^T\right)^{-1}\mathbf{y}$$

[$\mathbf{K}$ is of dimension $m \times n$, so the inverted matrix $\mathbf{K}\mathbf{K}^T$ is of dimension $m \times m$, which is OK since $m < n$]

In component notation: associating the $m$ constraining equations ($y_i = \sum_j K_{ij}x_j$ for the exact solution) with a set of $m$ Lagrange multipliers $\lambda_i$:

$$\frac{\partial}{\partial x_k}\left(\sum_j x_j^2 + \sum_i \lambda_i\big(y_i - \sum_j K_{ij}x_j\big)\right) = 0 \quad \text{[NB: } \textstyle\sum_j K_{ij}x_j - y_i \text{ is also OK} \rightarrow \text{gives same solution]}$$

$$2x_k - \sum_i \lambda_i K_{ik} = 0$$

$$x_k = \frac{1}{2}\sum_i K_{ik}\lambda_i$$

Substituting $x_k$ into the contraint: $y_l = \sum_k K_{lk}x_k$:

$$\begin{aligned} y_l &= \frac{1}{2}\sum_k K_{lk}\sum_i K_{ik}\lambda_i \\ &= \frac{1}{2}\sum_i\sum_k K_{lk}K^T_{ki}\lambda_i \\ &= \frac{1}{2}\sum_i \left(\mathbf{K}\mathbf{K}^T\right)_{li}\lambda_i \end{aligned}$$

This is a matrix equation of the form $y_l = A_{li}\lambda_i$ where $\mathbf{A}$ can be inverted to give:

$$\lambda_i = 2\sum_l \left(\mathbf{K}\mathbf{K}^T\right)^{-1}_{il} y_l$$

Substituting this in the previous equation for $x_k$ gives:

$$\begin{aligned} x_k &= \sum_i K_{ik}\sum_l \left(\mathbf{K}\mathbf{K}^T\right)^{-1}_{il} y_l \\ &= \sum_i K^T_{ki}\sum_l \left(\mathbf{K}\mathbf{K}^T\right)^{-1}_{il} y_l \\ \mathbf{x} &= \mathbf{K}^T\left(\mathbf{K}\mathbf{K}^T\right)^{-1}\mathbf{y} \end{aligned}$$

[4]

*(b) The least squares solution for the overconstrained problem, i.e. the one for which $\sum_i \epsilon_i^2$ is minimised. ($\boldsymbol{\epsilon} = \mathbf{y} - \mathbf{K}\mathbf{x}$)*

In vector notation: the cost term is represented by $\sum_i \epsilon_i^2 = (\mathbf{y} - \mathbf{K}\mathbf{x})^T(\mathbf{y} - \mathbf{K}\mathbf{x})$. Minimising this with respect to the vector $\mathbf{x}$:

$$\frac{\partial}{\partial \mathbf{x}}\left((\mathbf{y} - \mathbf{K}\mathbf{x})^T(\mathbf{y} - \mathbf{K}\mathbf{x})\right) = 0$$
$$-2\mathbf{K}^T(\mathbf{y} - \mathbf{K}\mathbf{x}) = 0$$
$$-\mathbf{K}^T\mathbf{y} + \mathbf{K}^T\mathbf{K}\mathbf{x} = 0$$
$$\mathbf{x} = \left(\mathbf{K}^T\mathbf{K}\right)^{-1}\mathbf{K}^T\mathbf{y}$$

[$\mathbf{K}$ still of dimension $m \times n$, but inverted matrix $\mathbf{K}^T\mathbf{K}$ is of dimension $n \times n$, which is OK here since $n < m$]

In component notation [p.41 lecture notes]:

$$\frac{\partial}{\partial x_k}\left(\sum_i \big(y_i - \sum_j K_{ij}x_j\big)^2\right) = 0$$
$$2\sum_i K_{ik}\big(y_i - \sum_j K_{ij}x_j\big) = 0$$
$$\sum_j \sum_i K^T_{ki} K_{ij} x_j = \sum_i K^T_{ki} y_i$$
$$\left(\mathbf{K^T K}\right)\mathbf{x} = \mathbf{K}^T\mathbf{y}$$
$$\mathbf{x} = \left(\mathbf{K}^T\mathbf{K}\right)^{-1}\mathbf{K}^T\mathbf{y}$$

[4]

*(c) The solution that minimises a weighted sum $\sum_i \epsilon_i^2 + \lambda \sum_j x_j^2$.*

In vector notation [p.43 of lecture notes, although redefining $\lambda \to 1/\lambda$, so the 'constraint' with the Lagrange multiplier $\lambda$ is now regarded as $\sum_j x_j^2$ term rather than the $\sum_i \epsilon_i^2$ term]:

$$\frac{\partial}{\partial \mathbf{x}}\left(\lambda \mathbf{x}^T\mathbf{x} + (\mathbf{y} - \mathbf{K}\mathbf{x})^T(\mathbf{y} - \mathbf{K}\mathbf{x})\right) = 0$$
$$2\lambda\mathbf{x} - 2\mathbf{K}^T(\mathbf{y} - \mathbf{K}\mathbf{x}) = 0$$
$$\left(\lambda\mathbf{I}_n + \mathbf{K}^T\mathbf{K}\right)\mathbf{x} = \mathbf{K}^T\mathbf{y}$$
$$\mathbf{x} = \left(\mathbf{K}^T\mathbf{K} + \lambda\mathbf{I}_n\right)^{-1}\mathbf{K}^T\mathbf{y}$$

In component notation:

$$\frac{\partial}{\partial x_k}\left(\lambda\sum_j x_j^2 + \sum_i \big(y_i - \sum_j K_{ij}x_j\big)^2\right) = 0$$
$$2\lambda x_k - 2\sum_i K_{ik}\big(y_i - \sum_j K_{ij}x_j\big) = 0$$
$$\lambda x_k + \sum_j \sum_i K^T_{ki} K_{ij} x_j = \sum_i K^T_{ki} y_i$$
$$(\lambda\mathbf{I}_n + \mathbf{K^T K})\mathbf{x} = \mathbf{K}^T\mathbf{y}$$
$$\mathbf{x} = \left(\mathbf{K}^T\mathbf{K} + \lambda\mathbf{I}_n\right)^{-1}\mathbf{K}^T\mathbf{y}$$

[4]

*How would you choose $\lambda$ to provide a satisfactory solution? Why might this solution be preferable to both (a) and (b)?*

[p44 lecture notes] Choose $\lambda$ so that $(\mathbf{y} - \mathbf{K}\mathbf{x})^T(\mathbf{y} - \mathbf{K}\mathbf{x}) = m\sigma^2$ (where $\sigma^2$ is the measurement variance). Preferable to (a) and (b) since the extra constraint acts as a smoothing term, and guarantees that the matrix inversion is non-singular so the same equation can be applied for all $m > n$, $m = n$ and $m < n$.

[3]

Total: [15]

**RS.8**

*Show that* $(\mathbf{K}^T\mathbf{K}+\lambda\mathbf{I}_n)^{-1}\mathbf{K}^T = \mathbf{K}^T(\mathbf{K}\mathbf{K}^T+\lambda\mathbf{I}_m)^{-1}$ *where* $\mathbf{K}$ *is a general* $m \times n$ *matrix (note that the two unit matrices* $\mathbf{I}$ *are of different sizes, indicated by the subscripts).*

Start by factorising $\mathbf{K}^T\mathbf{K}\mathbf{K}^T + \lambda\mathbf{K}^T$ two different ways and equating results:

$$\mathbf{K}^T\left(\mathbf{K}\mathbf{K}^T+\lambda\mathbf{I}_m\right) = \left(\mathbf{K}^T\mathbf{K}+\lambda\mathbf{I}_n\right)\mathbf{K}^T$$

The presence of the $\lambda\mathbf{I}$ terms guarantees that both matrices in brackets are non-singular, so multiplying by their inverses gives the required result:

$$\left(\mathbf{K}^T\mathbf{K}+\lambda\mathbf{I}_n\right)^{-1}\mathbf{K}^T = \mathbf{K}^T\left(\mathbf{K}\mathbf{K}^T+\lambda\mathbf{I}_m\right)^{-1}$$

[5]

*Hence show that the solutions to 7(a) and 7(b) are particular cases of 7(c).*

The above (l.h.s.) is the solution to 7(c). The limit $\lambda \to 0$ effectively removes the $\sum_j x_j^2$ constraint (that the solution be smooth) leaving just the $\sum_i \epsilon_i^2$ minimisation, which defines the least-squares solution 7(b).

Also setting $\lambda \to 0$ for the the equivalent 7(c) solution (r.h.s.) clearly gives the solution 7(a) however this *not* due to removing the smoothness constraint, which is still part of 7(a). The distinction is best explained by first dividing the 7(c) cost function by $\lambda$:

$$\frac{1}{\lambda}\sum_i \epsilon_i^2 + \sum_j x_j^2$$

Now as $(1/\lambda) \to \infty$, the emphasis on reducing $\sum_i \epsilon_i^2$ increases until an exact fit $\sum_i \epsilon_i^2 = 0$ is forced, while still retaining the smoothness constraint, hence 7(a).

Which limit is reached as $\lambda \to 0$ is determined by $m < n$ ($\mathbf{K}\mathbf{K}^T$ invertible, so 7(a)) or $m > n$ ($\mathbf{K}^T\mathbf{K}$ invertible, so 7(b)).

[5]

Total: [10]

**RS.9**

*Limb darkening is the decrease of radiance* $L_\nu(\theta)$ *leaving the top of an atmosphere, with increase of angle* $\theta$ *to the zenith. Explain how limb darkening of thermal emission is related to the distribution of temperature in the atmosphere. Is it possible for an atmosphere to show limb darkening in one part of the spectrum and limb brightening in another?*

[p13 lecture notes]. For nadir viewing, the radiance is given by the multiplication of the Planck Function by a 'weighting function' $d\tau/dz$, and the brightness temperature of the radiance is largely determined by the atmospheric temperature at the peak of the weighting function.

Viewing at an angle to the vertical, the transmittance decreases more rapidly for any given atmospheric depth so the peak is shifted higher in altitude. Thus if the peak of the nadir weighting function occurs where the temperature is decreasing with altitude, there will be limb-darkening, and if the temperature increases with altitude there will be limb-brightening.

Different spectral regions have different weighting functions (according to the optical properties of the atmosphere as a function of wavelength) so it is possible to have weighting functions which peak in regions of positive or negative temperature gradients (e.g. troposphere-stratosphere, or stratosphere-mesosphere) and therefore possible to observe both limb-darkening and limb-brightening by observing the same atmosphere in different spectral regions.

[3]

*Assuming that the variation of the Planck function* $B_\nu(z)$ *with height* $z$ *in a plane parallel atmosphere can be expressed as a power series in optical depth* $\chi_\nu(z)$*:*

$$B_\nu(z) = \sum_{n=0}^{\infty} a_{\nu n}[\chi_\nu]^n$$

*find an expression for the radiance* $L_\nu(\theta)$ *emitted at zenith angle* $\theta$ *at the top of the atmosphere.*

Optical depth $\chi$ is measured in the nadir direction, so $\chi = 0$ is the top of the atmosphere. Optical depth is related to transmittance $\tau$ by $\tau = \exp(-\chi)$, so $d\tau/dz = -\exp(-\chi)d\chi/dz$ and the the radiance emitted vertically is given by:

$$L_\nu(0) = \int_0^\infty B_\nu(z)\exp(-\chi)\,d\chi$$

For an angle $\theta$, the path is modified by a factor $\sec\theta$:

$$L_\nu(\theta) = \int_0^\infty B_\nu(z)\exp(-\chi\sec\theta)\sec\theta\, d\chi$$

Substituting the given expression for $B(z)$ (and dropping the $\nu$ subscripts):

$$L(\theta) = \int_0^\infty \sum_{n=0}^{\infty} a_n\chi^n \exp(-\chi\sec\theta)\sec\theta\, d\chi$$

Changing to variable $y = \chi\sec\theta$, $\chi^n = y^n\cos^n\theta$:

$$L(\theta) = \sum_{n=0}^{\infty} a_n\cos^n\theta \int_0^\infty y^n\exp(-y)\, dy$$

Consider the integral equation:

$$I(n) = \int_0^\infty x^n e^{-x}\, dx$$

Integrating by parts:

$$I(n) = \left[-x^n e^{-x}\right]_0^\infty + n\int_0^\infty x^{n-1}e^{-x}\, dx = 0 + nI(n-1)$$

Since $I(0) = \int_0^\infty e^{-x}\, dx = 1$, this gives $I(n) = n!$. Hence:

$$L(\theta) = \sum_{n=0}^{\infty} a_n n!\cos^n\theta$$

[8]

*Describe a simple method (ignoring experimental error) of using measurements of limb darkening at one wavelength in order to derive the variation of the Planck function with $z$, and hence the temperature profile. Discuss the range of optical depth over which your method is likely to be reliable if the measurement of $L_\nu(\theta)$ is subject to experimental error. Comment briefly on the practicability of the measurement of limb darkening as a technique for remote temperature sounding.*

The coefficients $a_n$ can be determined from a series of measurements of $L(\theta)$ at known angles $\theta$. These are then used to derive $B(\chi)$ using the given power series [also need to assume that $\chi(z)$ is known in order to get $B(z)$ as suggested by the question]

The technique may be OK for determining a linear or quadratic form of $B(\chi)$ but higher order coefficients are more sensitive functions of the differences between measured radiances, hence their determination would be limited by noise. Also, since the technique relies on the vertical temperature profile being the same for all slant paths, any horizontal temperature gradients in the atmosphere will create further errors.

Information from deepest in atmosphere will come from nadir viewing, with a weighting function peaking at optical depth $\chi = 1$ (say). Information from highest in atmosphere comes from slant viewing, for which the plane-parallel atmosphere assumption will be valid up to (say) $\sec\theta = 10$ ($\theta = 84°$), equivalent to a nadir weighting function peaking at optical depth $\chi = 1/\sec\theta = 0.1$. Thus the vertical coverage is $\chi = 0.1$ to $\chi = 1$, which is poor. [the numbers used in this part are fairly arbitrary].

[4]

[Further considerations: assume that the optical depth varies as pressure: $\chi = \beta p \Rightarrow \tau = \exp(-\beta p)$ (see RS.4). From RS.5 (with $\alpha = 1$) the peak of the weighting function will occur at $p_{\max} = 1/\beta \Rightarrow \chi = 1$ (as assumed above) with a halfwidth of ~2.5 scale heights. At an optical depth of $\chi = 0.1$ (taken as the maximum altitude for which this technique is valid), $p = 0.1p_{\max}$, i.e. 2.3 scale heights above the the peak of the nadir weighting function. Thus the resolution (2.5 scale heights) is comparable with the vertical range (2.3 scale heights) so vertical resolution would be poor.]

Total: [15]

---

**Total for RS Sheet: 100**

[10.04]