

系外惑星と推定：東北大集中講義 2025 資料

河原創

2025 年 9 月 20 日

目次

第 1 章	Introduction	3
1.1	諸量 *	3
第 2 章	惑星の運動	4
2.1	二体問題 *	4
2.2	三次元空間における二体問題	6
2.3	二体問題を時間について求める	6
2.4	視線速度	8
2.5	アストロメトリ	10
第 3 章	運動の推定	12
3.1	点推定と最適化	12
3.2	最適化問題と自動微分	13
3.3	ベイズの定理	15
3.4	マルコフ鎖モンテカルロ	17
3.5	ハミルトニアン・モンテカルロと自動微分	18
3.6	ガウス過程 ‡	18
3.7	ガウス過程ノイズを含んだモデルフィット ‡	22
第 4 章	系外惑星からのスペクトル	24
4.1	トランジット現象	25
4.2	透過分光スペクトル	27
4.3	惑星からの光	27
4.4	輻射スペクトル	29
4.5	反射・散乱スペクトル	30
4.6	様々な系外惑星スペクトル	31
第 5 章	系外惑星大気	33
5.1	熱化学平衡における分子混合率	33
5.2	二原子分子の振動・回転遷移	37
5.3	ラインプロファイルと線強度	41
5.4	様々な分子の振動回転遷移 ‡	42
5.5	重力と惑星大気	44

5.6	大気と分子存在量	46
第 6 章	スペクトルからの大気推定	48
6.1	透過光スペクトルのモデリング	48
6.2	Schwarzschild 方程式 *	50
6.3	輻射光スペクトルのモデリング	51
6.4	散乱・反射スペクトルのモデリング (二流近似) [‡]	52
参考文献		59

第 1 章

Introduction

各章に * がついているものは学部の復習であるので授業では簡単に復習するにとどめる。† がついている部分は、細かい話や計算の導出である。基本、飛ばすので暇な人は読んでください。♣ は計算のヒント。‡ がついているものは本題から外れ気味な部分や、プラスアルファの部分。

1.1 諸量 *

天体の記号は \odot : 太陽、 J : 木星、 \oplus : 地球を表す。 R_x は天体 x の半径、 M_x は天体 x の質量を表す。

長さや角度

各天体の半径はだいたい

- $R_{\odot} = 7 \times 10^5 \text{ km}$
- $R_J = R_{\odot}/10$
- $R_{\oplus} = R_J/10 = R_{\odot}/100$

くらいである。太陽のシュワルツシルド半径

- $r_g = 2GM_{\odot}/c^2 = 3 \text{ km}$

くらいだが、以下の便利な関係を使うとよい。

- $\frac{r_g}{2 \text{ au}} = 10^{-8}$

太陽の直径は 1 au の 1/100 である。

- $\frac{2R_{\odot}}{1 \text{ au}} = 10^{-2}$

一番近い星までの距離は 1 pc くらい。1'' という角度について：

- $1'' = 5 \times 10^{-6} \text{ radian}$
- $\frac{1 \text{ au}}{1 \text{ pc}} = 1''$

- 典型的なシーイング $\sim 1''$
- 月の視直径 $\sim 30'$

重さ

- $M_{\odot} = 2 \times 10^{33} \text{ g}$
- $M_J = M_{\odot}/1000$
- $M_{\oplus} = M_J/300$

第2章

惑星の運動

系外惑星の観測量として惑星の運動に関するものを扱う。運動と関係する観測量としては、惑星による恒星の視線速度の運動である視線速度カーブ、天球方向の運動であるアストロメトリの二つがある。

2.1 二体問題*

惑星と恒星の運動と軌道を考える。まず惑星と恒星をそれぞれ質点と考え(質量 m_1, m_2 、座標 $\mathbf{r}_1, \mathbf{r}_2$)、惑星と恒星の間に働く力を重力だけと考えることで、二体問題となり解析的に解くことができる。図2.1のように $\mathbf{r} \equiv \mathbf{r}_2 - \mathbf{r}_1$ とおく。軌道面上の円座標系 $\mathbf{e}_r = (\cos \theta, \sin \theta)^\top$ と $\mathbf{e}_\theta = (-\sin \theta, \cos \theta)^\top$ で考える。 $\mathbf{r} = r\mathbf{e}_r$ であり、また $\dot{\mathbf{e}}_r = \dot{\theta}(-\sin \theta, \cos \theta)^\top = \dot{\theta}\mathbf{e}_\theta$ と $\dot{\mathbf{e}}_\theta = \dot{\theta}(\cos \theta, \sin \theta)^\top = \dot{\theta}\mathbf{e}_r$ を用いると、

$$\dot{\mathbf{r}} = \frac{d}{dt} r\mathbf{e}_r = \dot{r}\mathbf{e}_r + r\dot{\mathbf{e}}_r = \dot{r}\mathbf{e}_r + r\dot{\theta}\mathbf{e}_\theta \quad (2.1)$$

$$\ddot{\mathbf{r}} = (\ddot{r} - r\dot{\theta}^2)\mathbf{e}_r + \left[\frac{1}{r} \frac{d}{dt}(r^2\dot{\theta}) \right] \mathbf{e}_\theta \quad (2.2)$$

となる。重心を原点にとった時、 $\mathbf{r}_1 = -m_2/(m_1 + m_2)\mathbf{r}$ 、 $\mathbf{r}_2 = m_1/(m_1 + m_2)\mathbf{r}$ であるから、ラグランジアンは、換算質量 $\mu = (m_1^{-1} + m_2^{-1})^{-1}$ を用いて、

$$\begin{aligned} L = T - U &= \frac{m_1}{2}|\dot{\mathbf{r}}_1|^2 + \frac{m_2}{2}|\dot{\mathbf{r}}_2|^2 + G\frac{m_1 m_2}{r} \\ &= \frac{\mu}{2}[\dot{r}^2 + (r\dot{\theta})^2] + G\frac{m_1 m_2}{r} \end{aligned} \quad (2.3)$$

となる。

問題

ラグランジュ方程式

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{r}} \right) - \frac{\partial L}{\partial r} = 0, \quad \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\theta}} \right) - \frac{\partial L}{\partial \theta} = 0 \quad (2.4)$$

から

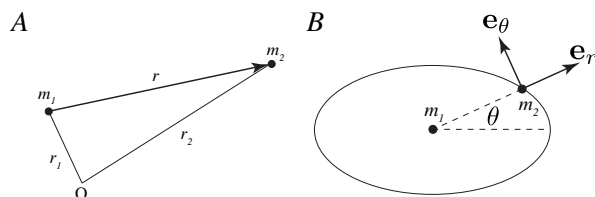


図2.1 座標系

$$\ddot{r} - r\dot{\theta}^2 = -\frac{G(m_1 + m_2)}{r^2} \quad (2.5)$$

と $\dot{\mathbf{h}} = 0$ が得られることを確かめよ。ただし $\mathbf{h} \equiv \mathbf{r} \times \dot{\mathbf{r}}$ は軌道角運動ベクトルである。
すなわち軌道角運動量ベクトル \mathbf{h} は、二体問題上は保存量である。これは質点が \mathbf{h} に直交する平面(軌道面)上の運動に制約されていることを意味する。

式(2.5)を変数変換し、 $u \equiv 1/r$ とする。 $hu^2 = \dot{\theta}$ 、 $\frac{dr}{dt} = \frac{dr}{d\theta} hu^2 = -h \frac{du}{d\theta}$ を用いて、

$$\frac{d^2 u}{d\theta^2} + u = \frac{G(m_1 + m_2)}{h^2} \quad (2.6)$$

となる。これは非斉次二階微分方程式であり、斉次解は

$$u_h = c_1 \cos \theta + c_2 \sin \theta = C_1 \cos(\theta - C_2) \quad (2.7)$$

のように書くことができる。非斉次解は明らかに

$$u_i = \frac{G(m_1 + m_2)}{h^2} \quad (2.8)$$

であり、これらを足しあわせたものが一般解である。 $G(m_1 + m_2)/h^2$ をくくり出し、積分定数項を e, ω とした一般解は

$$u = \frac{G(m_1 + m_2)}{h^2} [1 + e \cos(\theta - \omega)] \quad (2.9)$$

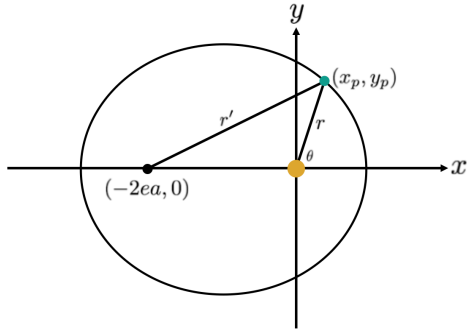


図2.2 楕円

とかける。 $r = 1/u$ に戻すと、

$$r = \frac{h^2}{G(m_1 + m_2)} \frac{1}{1 + e \cos(\theta - \omega)} \quad (2.10)$$

ところで、図2.3のような楕円を考えてみよう。楕円なので

$$r + r' = 2a \quad (2.11)$$

が成り立つ。また座標から

$$\begin{aligned} (r')^2 &= (x_p + 2ea)^2 + y_p^2 \\ &= (r \cos \theta + 2ea)^2 + (r \sin \theta)^2 \end{aligned} \quad (2.12)$$

となる。式(2.11, 2.12)から r' を消去すると、楕円の conic 方程式

$$r = \frac{a(1 - e^2)}{1 + e \cos \theta} \quad (2.13)$$

が得られる。

長半径 a 、短半径 b 、また true anomaly f を

$$\frac{h^2}{G(m_1 + m_2)} = a(1 - e^2) \quad (2.14)$$

$$b^2 = a^2(1 - e^2) \quad (2.15)$$

$$f \equiv \theta - \omega \quad (2.16)$$

と定義すれば、式(2.10)は conic 方程式の形

$$r = \frac{a(1 - e^2)}{1 + e \cos f} \quad (2.17)$$

になることから、2体問題の解が楕円になることが確認される(ただし $0 \leq e < 1$ とする)。また true

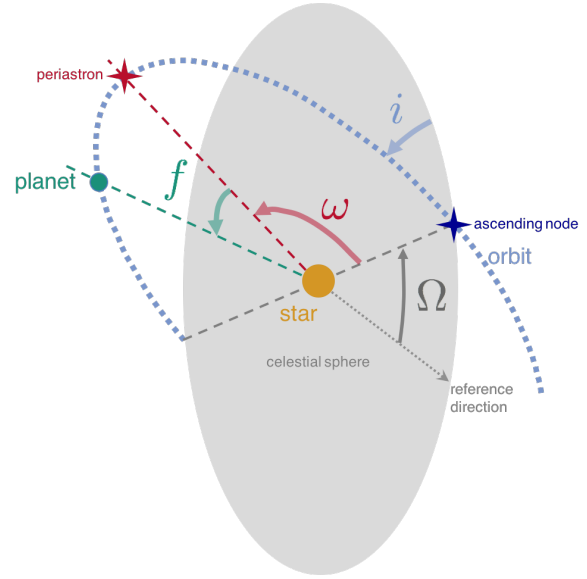


図2.3 座標系

anomaly f は、periastron (近日点の太陽を恒星に変えたもの) から測った恒星惑星の角度に対応することが分かる。 ω は the argument of periastron と呼ばれる量で conic 方程式(2.13)で表される楕円を半時計回りに ω 回転したということを表している。

ところで \dot{A} は一定であるから楕円面積

$$A = \pi ab = \frac{h}{2}P \quad (2.18)$$

とかける。ここに P は公転周期である。これを変形すると

$$P^2 = \frac{4\pi^2}{G(m_1 + m_2)} a^3 \quad (2.19)$$

となる。公転周期の自乗は質量に反比例、軌道長半径の3乗に比例することが分かる(ケプラーの第三法則)。円軌道の場合、公転角速度は $1/\sqrt{G(m_1 + m_2)}a$ であることも分かる。式(2.14)を用いて G を消すことで角運動量表記

$$P = \frac{2\pi a^2 \sqrt{1 - e^2}}{h} \quad (2.20)$$

もえられる。

平均運動は

$$n \equiv \frac{2\pi}{P} \quad (2.21)$$

で定義される。式 (2.20) を用いることで

$$n = \frac{h}{a^2 \sqrt{1-e^2}} \quad (2.22)$$

とも表せる。

2.2 三次元空間における二体問題

前章では、二次元平面上での二体運動を考えた。実際の天体は観測者から見て三次元空間上に存在しているので、軌道を3次元回転する必要がある。観測者からみた傾きである軌道傾斜角 (orbital inclination) を i で、azimuth 方向の定義を longitude of ascending node Ω で指定する。これら $a, e, \omega, f, i, \Omega$ の六つのパラメタを指定すれば、三次元空間における二体問題の軌道が一意に決定される。

これらの回転を回転行列で表してみよう。まず、元となる楕円は conic 方程式 (2.13) であらわされる図2.3のものとしよう。 z 軸を紙面に垂直にとるとする。

- (1) conic 方程式 (2.13) を z 軸周りに反時計回りに ω 回したものが二体問題の楕円の方程式 (2.17) であった。
- (2) 次に x 軸周りに反時計方向に i 回すと軌道面が天球に対して i 傾く事になる。
- (3) 最後に z 軸周りに Ω 、つまり天球の azimuth 方向の回転の自由度が残る。

これらの回転行列を順に楕円を表す $\mathbf{r} = (r \cos \theta, r \sin \theta, 0)$ に対してかける。まず (1) は

$$\mathbf{r} = \begin{pmatrix} r \cos \theta \\ r \sin \theta \\ 0 \end{pmatrix} = \begin{pmatrix} r \cos(f + \omega) \\ r \sin(f + \omega) \\ 0 \end{pmatrix} \quad (2.23)$$

である。次に (2) は

$$\mathbf{r}' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos i & -\sin i \\ 0 & \sin i & \cos i \end{pmatrix} \mathbf{r} \quad (2.24)$$

$$= \begin{pmatrix} r \cos(f + \omega) \\ r \sin(f + \omega) \cos i \\ r \sin(f + \omega) \sin i \end{pmatrix} \quad (2.25)$$

となる。最後に (3) は

$$\mathbf{r}'' = \begin{pmatrix} \cos \Omega & -\sin \Omega & 0 \\ \sin \Omega & \cos \Omega & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{r}' \quad (2.26)$$

$$= \begin{pmatrix} r \cos(f + \omega) \cos \Omega - r \sin(f + \omega) \cos i \sin \Omega \\ r \cos(f + \omega) \sin \Omega + r \sin(f + \omega) \cos i \cos \Omega \\ r \sin(f + \omega) \sin i \end{pmatrix} \quad (2.27)$$

$$\equiv \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}. \quad (2.28)$$

となり、三次元空間での軌道 (X, Y, Z) が得られた。

2.3 二体問題を時間について求める

二体問題の解は true anomaly f の関数、つまり $r = r(f)$ として書かれていた。しかし現実の観測では、二体問題はなんであれ時間の関数として観測される。そこで時間表記 $r = r(t)$ をどのように求めるか考えてみよう。

方針は以下のとおりである。まず

$$v^2 = \dot{\mathbf{r}} \cdot \dot{\mathbf{r}} = (\dot{r} \mathbf{e}_r + r \dot{f} \mathbf{e}_\theta) \cdot (\dot{r} \mathbf{e}_r + r \dot{f} \mathbf{e}_\theta) \quad (2.29)$$

$$= \dot{r}^2 + r^2 \dot{f}^2 \quad (2.30)$$

$$= \dot{r}^2 + \frac{h^2}{r^2}$$

を考える ($h = r^2 \dot{f}$ を用いた)。そして、

- (1) v^2 を式 (2.30) から r, \dot{r} を f の関数に変換し $v^2(f)$ で表す
- (2) $v^2(f)$ を Conic 方程式で $v^2(r)$ に変換
- (3) $v^2 = \dot{r}^2 + h^2/r^2$ と等置して r, \dot{r} の式、つまり r の時間に関する微分方程式を得る

(1) について r から f への変換は Conic 方程式 (2.17) を用いればよい。また \dot{r} を f に変換するには

$$\dot{r} = \frac{h}{a(1-e^2)} e \sin f \quad (2.31)$$

を用いる。

♣ 式 (2.31) の導出

$$\begin{aligned}\dot{r} &= \frac{df}{dt} \frac{d}{df} \left(\frac{a(1-e^2)}{1+e\cos f} \right) = \dot{f} \frac{a(1-e^2)e\sin f}{(1+e\cos f)^2} \\ &= \dot{f} \frac{a(1-e^2)}{1+e\cos f} \frac{e\sin f}{1+e\cos f} = r \dot{f} \frac{e\sin f}{1+e\cos f} \\ &= \frac{h}{r} \frac{e\sin f}{1+e\cos f} = \frac{h}{a(1-e^2)} e\sin f \quad (2.32)\end{aligned}$$

つまり

$$\begin{aligned}v^2 &= \dot{r}^2 + (r\dot{f})^2 = \dot{r}^2 + \frac{h^2}{r^2} \\ &= \frac{h^2}{a^2(1-e^2)^2} [2(1+e\cos f) + e^2 - 1] = v^2(f) \quad (2.33)\end{aligned}$$

次に $v^2(f)$ を再度 Conic 方程式を用いて r の関数に変換する (2) の手順は、

$$v^2(r) = \frac{h^2}{a(1-e^2)} \left(\frac{2}{r} - \frac{1}{a} \right) \quad (2.34)$$

となる。(3) より r の時間微分方程式

$$\dot{r}^2 - \frac{h^2}{a(1-e^2)} \left(\frac{2}{r} - \frac{1}{a} \right) + \frac{h^2}{r^2} = 0 \quad (2.35)$$

を得る。この方程式はそのまま解くことはできない。そこで eccentric anomaly E を導入する。

$$r = a(1 - e\cos E) \quad (2.36)$$

この E を用いて、微分方程式 (2.35) を E, \dot{E} の微分方程式に変換すると

$$\dot{E} = \frac{n}{1 - e\cos E} \quad (2.37)$$

また、天下りであるが、この解は

$$E - e\sin E = n(t - t_0) \quad (2.38)$$

で与えられる (微分することで式 (2.37) が成り立つことを確かめよ)。ここで時間変数の代理として Mean anomaly M を

$$M \equiv n(t - t_0) \quad (2.39)$$

と定義することにより

$$f(E) = E - e\sin E - M = 0 \quad (2.40)$$

を解くことで E が求まる。

♣ 式 (2.37) の導出

$$\begin{aligned}\frac{2}{r} - \frac{1}{a} &= \frac{1}{a} \left(\frac{2}{1 - e\cos E} - 1 \right) \\ &= \frac{1}{a} \frac{1 + e\cos E}{1 - e\cos E} = \frac{1}{a} \frac{1 - e^2 \cos^2 E}{(1 - e\cos E)^2} \quad (2.41)\end{aligned}$$

$$\begin{aligned}\frac{h^2}{r^2} - \frac{h^2}{a^2(1-e^2)} \left(\frac{2}{r} - \frac{1}{a} \right) &= \frac{h^2}{a^2(1-e\cos E)^2} - \frac{h^2(1-e^2)}{a^2(1-e\cos E)^2(1-e^2)} \\ &= \frac{h^2(1-e^2)}{a^2(1-e\cos E)^2(1-e^2)} - \frac{h^2(1-e^2\cos^2 E)}{a^2(1-e\cos E)^2(1-e^2)} \\ &= -\frac{h^2 e^2 (1 - \cos^2 E)}{a^2(1-e\cos E)^2(1-e^2)} \quad (2.42)\end{aligned}$$

$$\dot{r} = \frac{dE}{dt} \frac{d}{dE} (a(1 - e\cos E)) = \dot{E} a e \sin E \quad (2.43)$$

$$\dot{r}^2 = \dot{E}^2 a^2 e^2 \sin^2 E \quad (2.44)$$

これらの式より、

$$\dot{E}^2 a^2 e^2 (1 - \cos^2 E) = \frac{h^2 e^2 (1 - \cos^2 E)}{a^2(1-e\cos E)^2(1-e^2)} \quad (2.45)$$

つまり

$$\dot{E}^2 = \frac{h^2}{a^4(1-e\cos E)^2(1-e^2)} \quad (2.46)$$

$\dot{E} > 0$ に軸を取るとすると、

$$\dot{E} = \frac{h}{a^2 \sqrt{1-e^2} (1-e\cos E)} \quad (2.47)$$

$$= \frac{n}{1-e\cos E} \quad (2.48)$$

を得た。最後の変形は平均運動 (2.22) を用いた。

また E と f の関係は式 (2.36) と二体問題の円錐方程式 (2.17) から

$$\cos f = \frac{\cos E - e}{1 - e\cos E} \quad (2.49)$$

これにより、周期 P とオフセット t_0 が分かれば時間 t から mean anomaly M が計算でき、数値計算で式 (2.40) を解くことで E が、さらに式 (2.49) から true anomaly f を求めることができる。

Newton-Raphson 法により式 (2.40) を解く

非線形方程式 $f(E) = 0$ は数値的に求めることが必要である。このための方法の一つとして Newton-Raphson 法を紹介する。Newton-Raphson

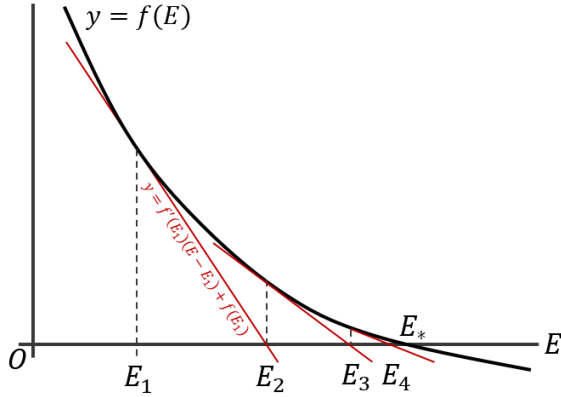


図2.4 Newton-Raphson 法 の概念図。

法は図2.4にしめすように、まず初期値 E_1 から初めて、 $f(E_1)$ に接する直線を求め、その直線と原点との交点を解析的に求め E_2 とする。この手続きを n 回繰り返すという手順である。 i 番目の手続きでは接線 (つまり $f(E)$ を E_i でテイラー展開した一次の項までの近似式) は

$$y = f(E_i) + f'(E_i)(E - E_i) \quad (2.50)$$

となることから

$$E_{i+1} = E_i + \Delta E_i \quad (2.51)$$

$$\Delta E_i \equiv -\frac{f(E_i)}{f'(E_i)} \quad (2.52)$$

となる。 ΔE_i を i 番目の更新項 (update term) とよぶ^{*1}。この手続きを収束条件の判定値を ϵ として $|E_{i+1} - E_i| < \epsilon$ となるまで繰り返し、条件を満たした E_i を近似解とすればよい。

式 (2.50) の計算には f の微分が必要である。式 (2.40) より

$$f'(E) = 1 - e \cos E \quad (2.54)$$

である。

^{*1} 単に $f(E)$ をテイラー展開して1次まで残した近似を用いて解を更新していくとみることでもできる。すなわち

$$f(E) \approx f(E_i) + f'(E_i)(E - E_i) = 0 \quad (2.53)$$

の解を E_{i+1} として、繰り返すことに対応している。

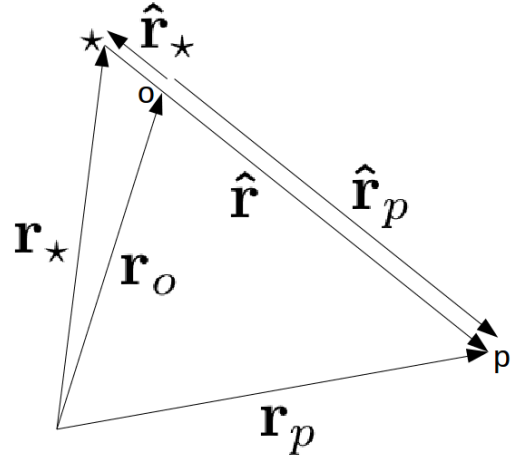


図2.5 重心 o と各ベクトルの定義。

二次最適化としての Newton 法[†]

コスト関数 $Q(x)$ を最小にする $x = x^*$ を求める最適化問題

$$x^* = \underset{x}{\text{minimize}} Q(x) \quad (2.55)$$

の解法として、停留点条件 $dQ(x)/dx = 0$ を Newton-Raphson 法で求めることができる。この場合、 $f(x) = Q'(x)$ として、式 (2.51) を用いると、更新式は

$$x_{i+1} = x_i + \Delta x_i \quad (2.56)$$

$$\Delta x_i \equiv -\frac{Q'(x_i)}{Q''(x_i)} \quad (2.57)$$

となり、 $Q(x)$ の二回微分が必要であることから、二次の最適化と呼ばれる。ちなみに最適化の文脈では Raphson が省略され、単に Newton 法と呼ばれることが多い。一般に多次元の Newton 法も同様に構成できる。多次元のテイラー展開より

$$f(x) \approx f(x_i) + J(x_i)(x - x_i) = 0 \quad (2.58)$$

をみたす x が次の更新点となるので、

$$x_{i+1} = x_i + \Delta x_i \quad (2.59)$$

$$\Delta x_i \equiv -J^{-1}(x_i)f(x_i) \quad (2.60)$$

となる。ただし $J(x_i)$ はヤコビアンである。

2.4 視線速度

太陽系外惑星は、惑星の公転による恒星の視線速度方向の変化を、恒星スペクトルのドップラーシフトを通じて検出することで初検出された。

惑星による恒星の視線速度変動を考える。二体の場合、恒星の視線速度は恒星と惑星の重心から測つ

たベクトルの運動の視線速度が観測される。そこで、図2.5のように惑星 p 、恒星 \star とその重心 o をおく。主星、惑星質量をそれぞれ M_\star 、 M_p とし、主星と惑星の位置を \mathbf{r}_\star 、 \mathbf{r}_p とする。この系の重心 \mathbf{r}_o は

$$\mathbf{r}_o = \frac{M_\star \mathbf{r}_\star + M_p \mathbf{r}_p}{M_\star + M_p} \quad (2.61)$$

である。重心から測った主星と惑星の位置 $\hat{\mathbf{r}}_\star = \mathbf{r}_\star - \mathbf{r}_o$ 、 $\hat{\mathbf{r}}_p = \mathbf{r}_p - \mathbf{r}_o$ 、また主星から惑星に向かうベクトルを $\hat{\mathbf{r}} = \hat{\mathbf{r}}_p - \hat{\mathbf{r}}_\star$ と定義する。

$$M_\star \hat{\mathbf{r}}_\star + M_p \hat{\mathbf{r}}_p = M_\star \hat{\mathbf{r}}_\star + M_p (\hat{\mathbf{r}}_\star + \hat{\mathbf{r}}) = 0 \quad (2.62)$$

であるから

$$\hat{\mathbf{r}}_\star = -\frac{M_p}{M_\star + M_p} \hat{\mathbf{r}} \quad (2.63)$$

である。

まず、観測者から恒星への位置ベクトルは重心までの位置ベクトル \mathbf{r}_o を用いて

$$\mathbf{r}_\star = \mathbf{r}_o + \hat{\mathbf{r}}_\star \quad (2.64)$$

である。恒星は惑星と反対に動くので、視線方向を Z 軸の反対向きにとり、半時計回りを保つために $\Omega = \pi$ ととるとしよう。この場合、恒星の視線速度はこの時間微分の Z 方向への単位ベクトル \mathbf{e}_Z と内積をとったものを符号反転させたものだから、

$$v_r = V_{\text{sys}} - \dot{\mathbf{r}}_\star \cdot \mathbf{e}_Z \quad (2.65)$$

ここに $V_{\text{sys}} \equiv -\dot{\mathbf{r}}_o \cdot \mathbf{e}_Z$ は系全体の視線速度である。式 (2.63) より、

$$v_r = V_{\text{sys}} + \frac{M_p}{M_\star + M_p} \dot{\mathbf{r}} \cdot \mathbf{e}_Z \quad (2.66)$$

と変形できる。この $\hat{\mathbf{r}}$ は、二体問題の \mathbf{r} に対応させれば、そのまま定式化を用いる事ができる。すると $\dot{\mathbf{r}}_\star \cdot \mathbf{e}_Z$ は、三次元座標系の Z 成分の時間微分 \dot{Z} に対応する事になる。すなわち式 (2.28) を用いて

$$v_r = V_{\text{sys}} + \frac{M_p}{M_\star + M_p} \dot{Z} \quad (2.67)$$

となる。

$$\begin{aligned} \dot{Z} &= \frac{d}{dt} [r \sin i \sin (f + \omega)] \\ &= \dot{r} \sin i \sin (f + \omega) + r \dot{f} \sin i \cos (f + \omega) \end{aligned} \quad (2.68)$$

である。

式 (2.31) と $r \dot{f} = h/r$ と Conic 方程式 (2.17) をもちいて、

$$v_r = V_{\text{sys}} + \frac{M_p}{M_\star + M_p} \frac{h \sin i}{a(1-e^2)} [e \sin f \sin (f + \omega) + e \cos f \cos (f + \omega) + \cos (f + \omega)] \quad (2.69)$$

$$= V_{\text{sys}} + \frac{M_p}{M_\star + M_p} \frac{h \sin i}{a(1-e^2)} [\cos (f + \omega) + e \cos \omega] \quad (2.70)$$

となる。

♣			
$\sin f \sin (f + \omega)$	+	$\cos f \cos (f + \omega)$	=
$\cos (-f) \cos (f + \omega)$	-	$\sin (-f) \sin (f + \omega)$	=
$\cos (-f + f + \omega) = \cos \omega$			

もしくは h を陽に書き下すと

$$v_r = V_{\text{sys}} + K_\star [\cos (f + \omega) + e \cos \omega] \quad (2.71)$$

$$K_\star \equiv \frac{M_p \sin i}{\sqrt{1-e^2}} \sqrt{\frac{G}{(M_p + M_\star)a}} \quad (2.72)$$

が二体問題の視線速度カーブとなる。また視線速度変動のオーダーは

$$K_\star \sim 30 \text{m/s} \frac{M_p \sin i}{M_J} \left(\frac{M_\star}{M_\odot} \right)^{-1/2} \left(\frac{a}{\text{au}} \right)^{-1/2} \quad (2.73)$$

$$= 130 \text{m/s} \frac{M_p \sin i}{M_\oplus} \left(\frac{M_\star}{M_\odot} \right)^{-1/2} \left(\frac{a}{0.05 \text{au}} \right)^{-1/2} \text{ (HJ)}$$

$$= 0.1 \text{m/s} \frac{M_p \sin i}{M_\oplus} \left(\frac{M_\star}{M_\odot} \right)^{-1/2} \left(\frac{a}{\text{au}} \right)^{-1/2} \text{ (Earth)}$$

となる。また、この視線速度変動に対応するドップラーシフトは

$$\begin{aligned} \frac{\Delta \lambda}{\lambda} &\sim \frac{K_\star}{c} \\ &= 10^{-7} \frac{M_p \sin i}{M_J} \left(\frac{M_\star}{M_\odot} \right)^{-1/2} \left(\frac{a}{\text{au}} \right)^{-1/2} \end{aligned} \quad (2.74)$$

となる。現在の高分散分光器の分解能は $R \sim 10^5$ であるので、多数の分子吸収線（と光子数）をもちいて S/N をブーストすることが重要となる。またプラク

ティカルには、ヨードセルや周波数コムを用いた波長較正の安定化が重要である。

$$\begin{aligned}
K_{\star} &\sim \frac{M_p \sin i}{M_{\odot}} \left(\frac{M_{\star}}{M_{\odot}} \right)^{-1/2} \sqrt{\frac{GM_{\odot}}{a}} \\
&= \frac{M_p \sin i}{M_{\odot}} \left(\frac{M_{\star}}{M_{\odot}} \right)^{-1/2} \sqrt{\frac{2GM_{\odot}}{c^2} \frac{1}{2a}} c \\
&= \frac{M_J}{M_{\odot}} \frac{M_p \sin i}{M_J} \left(\frac{M_{\star}}{M_{\odot}} \right)^{-1/2} \sqrt{\frac{r_g}{2au} \frac{au}{a}} c \\
&= 10^{-7} c \frac{M_p \sin i}{M_J} \left(\frac{M_{\star}}{M_{\odot}} \right)^{-1/2} \left(\frac{a}{au} \right)^{-1/2} \\
&= 30 \text{ m/s} \frac{M_p \sin i}{M_J} \left(\frac{M_{\star}}{M_{\odot}} \right)^{-1/2} \left(\frac{a}{au} \right)^{-1/2}
\end{aligned}$$

時間の関数としての視線速度カーブ

さて、実際の観測では時間の関数として測定値が得られるので、この解を時間の関数で書きたい。時間 t から周期が分かると mean anomaly M かわかる。 M から eccentric anomaly E は式 (2.40) を数值的に解く。これらにより、式 (2.71) を数值的に時間で書きなおす事ができる。このように視線速度カーブから求まる物理量は式 (2.71) から $V_{\text{sys}}, K_{\star}, e, \omega$ である。また f に関係して周期 P と位相量（時刻のオフセット）も推定するパラメータとなる。図2.6は幾つかの視線速度カーブと対応する楕円軌道を書いたものである。ケプラー第二法則より天体が近点付近に来了時に急速に視線速度カーブが変化するのをイメージできるだろう。

Binary Mass Function[†]

そもそも視線速度解析は恒星惑星系以前に連星系で発展してきた。この場合、 $M_p \ll M_{\star}$ は成り立たない。式 (2.72) を、 $\star \rightarrow 1$ と $p \rightarrow 2$ で書き直し、ケプラー第三法則 (2.19) を用いて、右辺に観測量だけ、左辺に物理パラメータとわけて書くと

$$f \equiv \frac{M_2^3}{(M_1 + M_2)^2} \sin^3 i = \frac{PK_1^3}{2\pi G} (1 - e^2)^{3/2} \quad (2.75)$$

のようになる。この f が、一般の質量比の場合、星1の視線速度カーブの観測量 K_1 、 e 、 P のみからわかる量である。この f を binary mass function と呼ぶ。

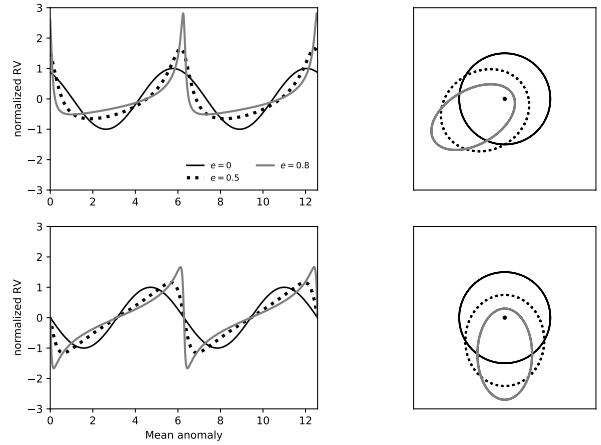


図2.6 視線速度カーブ (左) と対応する軌道 (右)。上段は $\omega = \pi/6$, 下段は $\omega = \pi/2$ である。線種は黒実線が $e = 0$ 、黒点線が $e = 0.5$ 、灰色実線が $e = 0.8$ に対応している。右の楕円は $i = \pi/2$ の時に軌道を上から見たものに対応し、下から上むきに視線方向をとると、視線速度カーブの恒星軌道、上から下向きにとると惑星の軌道に対応する。

また、逆に

$$\begin{aligned}
K_1 &= 29.8 \text{ [km/s]} \frac{\sin i}{\sqrt{1 - e^2}} \left(\frac{M_2}{M_1 + M_2} \right) \\
&\times \left(\frac{M_1 + M_2}{M_{\odot}} \right)^{1/3} \left(\frac{P}{1 \text{ yr}} \right)^{1/3} \quad (2.76)
\end{aligned}$$

とスケーリングされた式を用いると、連星系の場合の検出可能性の見積もりに便利である。この 29.8 km/s という値は地球の公転速度に一致している。

2.5 アストロメトリ

視線速度変動は3次元ケプラー運動の \dot{Z} 成分の運動情報を用いていた。アストロメトリは天球上の星の位置を計測する手法であるから、 (X, Y) 成分の運動情報を用いることになる。固有運動・パララックスを補正後の二体運動による恒星の位置は、観測者から恒星までの距離を d として、観測量である角度（離角）としてあらわすと

$$\theta = \frac{\hat{r}_{\star}}{d} = -\frac{M_p}{M_{\star} + M_p} \frac{\hat{r}}{d} \quad (2.77)$$

に二次元射影となる。すなわち式 (2.28) より、

$$\begin{aligned}\theta_x = & -\frac{M_p}{M_\star + M_p} \frac{X}{d} = -\frac{M_p}{M_\star + M_p} \frac{r}{d} \\ & \times [\cos(f + \omega) \cos \Omega - \sin(f + \omega) \cos i \sin \Omega]\end{aligned}\quad (2.78)$$

$$\begin{aligned}\theta_y = & -\frac{M_p}{M_\star + M_p} \frac{Y}{d} = -\frac{M_p}{M_\star + M_p} \frac{r}{d} \\ & \times [\cos(f + \omega) \sin \Omega + \sin(f + \omega) \cos i \cos \Omega]\end{aligned}\quad (2.79)$$

となる。この解は $(i \rightarrow \pi - i, f \rightarrow 2\pi - f, \omega \rightarrow 2\pi - \omega)$ の変換に不変であり、これはアストロメトリのみでは惑星が近づいているのか遠ざかっているのかはわからない。

第 3 章

運動の推定

第2章では系外惑星の運動に関する観測量の物理モデルを扱った。これらはすべて1次元の時系列データである。本章では一次元の時系列データから、物理モデルと統計モデルを組み合わせることでの内部パラメタである系外惑星の物理量の推定を行う。観測データ \mathbf{d} と理論モデル \mathbf{f} との比較により以下のようなことが得られる。

1. ある理論モデル \mathbf{f} は観測データを説明できるモデルかどうか知る
2. ある理論モデル \mathbf{f} が正しいとしたとき、そのモデル内に含まれるパラメタ $\boldsymbol{\theta}$ をデータ \mathbf{d} にもとづいて決定・推定する (パラメタ推定)
3. 複数のモデルが存在するときにどちらのモデルが観測 \mathbf{d} をよく説明できるか (モデル選択)

モデル選択はいろいろ大変なので、ここではパラメタ推定について考える。

ここでは例として以下にあげる視線速度の観測データと理論モデルの比較を行う。この場合、 \mathbf{d} は rv データ列を並べたベクトルと考えればよく、各要素 d_i はこれに対応する時間 t_i を通じて、理論モデルと比較可能となる。理論モデルは式 (2.71) である。パラメタベクトルは $\boldsymbol{\theta} = (V_{\text{sys}}, K_*, e, \omega, T_0, P)$ であることがわかる。

視線速度の生データ \mathbf{d}

<https://github.com/HajimeKawahara/class25>
を参照

3.1 点推定と最適化

パラメタ $\boldsymbol{\theta}$ を内部に持つモデル $\mathbf{f}(\boldsymbol{\theta})$ とデータ \mathbf{d} を何らかの手段で比較することで、一つのパラメタを選択するパラメタ推定を点推定という。もし、データに誤差が全く含まれていなかったら

$$\mathbf{d} - \mathbf{f}(\boldsymbol{\theta}) = \mathbf{0} \quad (3.1)$$

となる $\boldsymbol{\theta}$ を探せばよいだろう。しかし、通常、データには観測誤差が含まれていて、このような式を満たす $\boldsymbol{\theta}$ は通常見つからない。すなわち

$$\mathbf{d} = \mathbf{f}(\boldsymbol{\theta}^*) + \boldsymbol{\epsilon} \quad (3.2)$$

のようにかける。この式は、モデル \mathbf{f} が正しく、かつ”正しい”パラメタ $\boldsymbol{\theta}^*$ が存在し、残差ベクトル $\boldsymbol{\epsilon}$ が観測誤差のみに由来しているということを含意している。観測誤差を扱う形で理論構成を行うには、観測誤差を生成する確率モデルが必要であることがわかる。たとえば、 $\boldsymbol{\epsilon}$ が独立のゼロ平均の正規分布に従うとき、

$$\epsilon_i \sim \mathcal{N}(0, \sigma) \quad (3.3)$$

とかける。もしくはもっと簡潔に物理モデル $\mathbf{f}(\boldsymbol{\theta}^*)$ は確率モデルの平均値であるとして

$$d_i \sim \mathcal{N}(f_i(\boldsymbol{\theta}^*), \sigma) \quad (3.4)$$

という理論モデル構成を行うことができる。

さてこの議論でわかるのは、観測データと理論モデルを比較する、といったときの理論モデルとは、 \mathbf{f} だけではなく確率モデルも指定しないとならない。図式的に書くと

実証可能な理論モデル

実証可能な理論 = 物理・化学モデル + 確率モデル

通常、理論天文学では物理・化学モデルを重視しすぎるあまり確率モデルをおろそかにしがちである。また観測天文学では物理・化学モデルを固定しがちである。確率モデルの構成は、実際的な問題として計算量が多く、近年までは簡単には扱えなかった。しかし、最近では計算機の進歩とベイズ統計手法の進歩でこれに対処できるようになってきた。

さて、点推定に戻る。データ間の相関誤差がある場合も多いが、これはガウス過程によるモデリングに譲るとして、今は独立に誤差が入っているとしよう。さらにその誤差がどのデータ点に対しても同じ σ のガウシアンだとする。この場合、 θ をもつモデル $f(\theta)$ が d を生じさせる確率は

$$p(d|\theta) = \prod_{i=1}^N \mathcal{N}(f_i(\theta), \sigma) \quad (3.5)$$

$$\propto \exp\left(-\frac{\sum_{i=1}^N (f_i(\theta) - d_i)^2}{2\sigma^2}\right) \quad (3.6)$$

となる。これは尤度関数 (likelihood) である。尤度を最も大きくする θ を点推定値として採用する方法を最尤法という。この場合、

$$\theta^* = \text{minimize}_{\theta} L_2(\theta) \quad (3.7)$$

$$L_2(\theta) = \sum_{i=1}^N (f_i(\theta) - d_i)^2 \quad (3.8)$$

という最適化問題をとけばよいということになる。これは最小二乗法そのものである。すなわち最小二乗法は、誤差が独立なガウシアンであると仮定したときの最尤推定と一致する。

最小二乗法による点推定の仮定をもう少しゆるくしたものが、 χ^2 最小化である。これは、各データ点の σ_i が既知だとして

$$p(d|\theta) \propto \exp\left(-\sum_{i=1}^N \frac{(f_i(\theta) - d_i)^2}{2\sigma_i^2}\right) \quad (3.9)$$

となることより、

$$\theta^* = \text{minimize}_{\theta} \chi^2(\theta) \quad (3.10)$$

$$\chi^2(\theta) = \sum_{i=1}^N \frac{(f_i(\theta) - d_i)^2}{2\sigma_i^2} \quad (3.11)$$

の最適化問題を解くというものである。

3.2 最適化問題と自動微分

さてここまでの議論から、点推定の実実際的な問題系は最適化問題であるということがわかる。非線形モデルの最適化問題をとくには、以前見た二次最適化としての Newton 法や、その発展である準ニュートン法、諸先生方が Numerical Recipes を熟読しながら実装したマルカート法、もしくはアマーバー法としてしられる勾配を用いない Nelder-Mead 法など多種多様な方法がある。

しかしここでは、自動微分を用いた一次の最適化を考える。これはパラメタ数が多いと二次最適化はヘッシアン計算に律速されるためである。そのためニューラルネットなどの機械学習モデルは一次最適化を用いることが多い。一次最適化の基本是最急降下法であり

$$\theta^{(k)} = \theta^{(k-1)} - \gamma \frac{\partial}{\partial \theta} f(\theta) \quad (3.12)$$

のように、最も急な坂を下っていく方法である。最急降下法だと行き過ぎることがおおいので、さまざまなモーメントをつけて落ちやすくする。代表的な方法は ADAM[9] であるがここではアルゴリズムの詳細には踏み入れない。

一次最適化では、モデルのパラメタによる微分が必要である。数値的に微分値を求めるにはおおまかに 4 つの方法がある。一つ目は手で微分 (manual differentiation) をし、結果をコーディングすることである。これはモデルが複雑になってくると破綻しがちであり、またフレキシビリティの観点からもモデルの継続的な改良を妨げる。次に mathematica 等による symbolic differentiation を用いて手で微分する代わりに微分結果を得てコーディングすることが考えられる。これは mathematica 等を用いたことのある方ならわかると思うが、モデルが複雑になってくると膨大な項数の結果が排出されるので、やはりフレキシビ

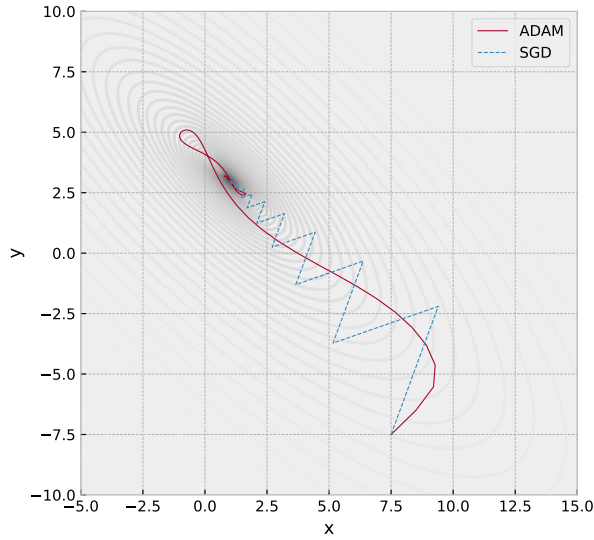


図3.1 一次最適化の例。

リティの観点からは難点がある。次に、数値微分を行うことが考えられる。数値微分はモデルが複雑になってくるとエラーがたまりやすい。そこで機械学習分野などで用いられているのが JAX/tensorflow/pytorch など用いられている、自動微分である。コーディングできるモデルは通常、様々な微分の既知である関数（ここでは要素関数と呼ぼう）の加減乗除の組み合わせでできている。そこで、自動微分では、連鎖則が成り立つように各要素関数

$$x \mapsto f(x) \quad (3.13)$$

を関数のアウトプットとその微分値の情報を出力するように拡張する。また微分の加乗除の演算規則がなりたつように演算を定義する。自動微分の実装法としては JAX では Jacobian Vector Product (JVP) を用いたものが使用されている。しかしここではイントロダクションとしてより説明が容易な、双対数を用いた方式を使用して自動微分を解説しよう。

双対数 (dual number)、 $z \in k[\epsilon]/\langle \epsilon^2 \rangle$ ^{*1} は $a, b \in \mathbb{R}$

^{*1} 多項式環の商環のことをしめしている。詳しくはたとえば雪江明彦「環と体とガロア理論」などを参照。

にたいし、

$$z = a + b\epsilon \quad (3.14)$$

$$\epsilon^2 = 0 \quad (3.15)$$

となる数である。複素数は $i^2 = -1$ であったのが $\epsilon^2 = 0$ となったと考えればよい。変数 x の拡張として実部 a に x を、非実部 b に x' を割り当てると、 $z = f + f'\epsilon, w = g + g'\epsilon$ に対し、可算・乗数・除算はそれぞれ

$$z + w = (f + g) + (f' + g')\epsilon \quad (3.16)$$

$$zw = fg + (f'g + fg')\epsilon \quad (3.17)$$

$$z/w = \frac{f}{g} + \frac{f'g - fg'}{g^2}\epsilon \quad (3.18)$$

となる。これは実部に関して通常の、また非実部に関して微分の加乗除則と同じであることを示している。次に連鎖則を実現するには関数 $x \mapsto F(x)$ を

$$x + x'\epsilon \mapsto F(x) + F'(x)x'\epsilon \quad (3.19)$$

と拡張すればよい。ここで $F(x)$ を元の関数、双対数をもつ拡張された関数を $\hat{F}(x + x'\epsilon)$ と分けて表記する。すると式 (3.19) は

$$\hat{F}(x + x'\epsilon) = F(x) + F'(x)x'\epsilon \quad (3.20)$$

と書ける。さて連鎖則 $G(F(x))$ の拡張は

$$\begin{aligned} \hat{G}(\hat{F}(x + x'\epsilon)) &= \hat{G}(F(x) + F'(x)x'\epsilon) \\ &= G(F(x)) + G'(F(x))F'(x)x'\epsilon \end{aligned} \quad (3.21)$$

となり、実部は通常の合成 $(G(F(x)))$ が、非実部は連鎖則 $G'(F(x))F'(x)x' = \frac{dG}{dF} \frac{dF}{dx} x'$ が実現されているのが確認できる。

ミニマム自動微分実装

というわけで小さな自動微分を Python で実装してみよう。ここで解きたい問題を

$$F(x) = \log(\cos x \sin x) + \sin x \quad (3.23)$$

の $x = 1$ で微分 $F'(1)$ とする。さらにその合成関数

$$G(x) = F(F(x)) \quad (3.24)$$

の $x = 1$ での微分 $G'(1)$ も求めたい。式 (3.24) を手で微分するのは大変であるし、symbolic differentiation

の結果も長大である。しかし自動微分ならばかなりすっきりしたコードになる。

まず関数 (3.23) に存在する演算は可算と乗算であるので、式 (3.16)、(3.17) に対応した双対数の可乗算を定義する。

```
1 def mul(x, y):
2     a, b = x
3     c, d = y
4     return a*c, a*d + b*c
5
6 def add(x, y):
7     a, b = x
8     c, d = y
9     return a + c, b + d
```

ここに x, y はそれぞれ双対数である。次に関数 (3.23) に存在する要素関数は $\sin x$, $\cos x$, $\log x$ である。式 (3.20) を実装するには、双対数の実部と非実部をペア (a, b) の形式であらわすと、関数 $F(x)$ に対し

$$(x, dx) \rightarrow (F(x), F'(x)dx) \quad (3.25)$$

となるように入出力を与えればよいことがわかる。そこで

```
1 import numpy as np
2 def cos(x):
3     a, b = x
4     return np.cos(a), - np.sin(a)*b
5
6 def sin(x):
7     a, b = x
8     return np.sin(a), np.cos(a)*b
9
10 def log(x):
11     a, b = x
12     return np.log(a), b/a
```

これで終わりである。あとは

```
1 f = lambda x:
2 add(log(mul(cos(x), sin(x))), sin(x))
```

```
3 df = lambda x: f([x, 1.0])
4 df(1.0)
```

とすれば $(F(1), F'(1))$ が $(0.05324076815279066, -0.37501280285243144)$ と求まる。 $G'(x)$ も簡単に、

```
1 g = lambda x: f(f(x))
2 dg = lambda x: g([x, 1.0])
3 dg(1.0)
```

とすることで $(-2.8816056725768977, -7.391555094461485)$ と求まる。以上の例からわかるように自動微分の実際の計算には代数計算を用いているため数値微分に比べて誤差の蓄積は少ない。またフレキシブルにコーディングが可能である。

ここでは自分で実装してみたが、実際には JAX や pytorch, Enzyme.jl などの自動微分パッケージを利用することになるだろう。さらに自動微分は、次に示すベイズ統計的なパラメタ推定でも威力を発する。このようにコードをエンドトゥーエンドで微分可能にしておくプログラミングを微分可能プログラミングという [1]。

3.3 ベイズの定理

ベイズの定理 (Bayes' theorem) とは、事象 A が起こる条件のもとで、別の事象 B が観測される確率を示す定理である。これは条件付き確率に関する基本的な関係式として知られている。

ベイズの定理を数式で表すと、以下のようなのである。

$$p(A | B) = \frac{p(B | A)p(A)}{p(B)}.$$

ここで、

- $p(A | B)$ は「 B が起きたときに A が起こる確率」である。
- $p(B | A)$ は「 A が起こったときに B が起こる確率」である。
- $p(A)$ は「 A が起こる確率」である。
- $p(B)$ は「 B が起こる確率」である。

ベイズの定理は、事象の背後にある確率的関係を逆方向から考察する際に極めて有用である。すなわち、ある観測結果 B から原因候補 A の確率を更新する方法を示してくれる。特に統計的推測や機械学習の分野において、事前確率 (prior probability) を観測データから得られる情報によって事後確率 (posterior probability) へと更新する根幹となる理論として広く用いられる。

証明は以下のとおりである。まず、条件付き確率の定義から、

$$p(A | B) = \frac{p(A \cap B)}{p(B)}$$

が成り立つ。一方、 A が起きたときに B が起こる確率を示す条件付き確率の定義より、

$$p(A \cap B) = p(B | A)p(A)$$

となる。これを上式に代入すれば、

$$p(A | B) = \frac{p(B | A)p(A)}{p(B)}$$

が得られる。

天文学の推定のコンテキストでは、 A をモデルの内の推定するパラメタ θ で、 B を観測データ d とおくことで事後確率を

$$p(\theta | d) = \frac{p(d | \theta)p(\theta)}{p(d)}.$$

と表して用いることが多い。すなわち $p(d | \theta)$ は点推定の時にも表れた尤度であり、 $p(\theta)$ はパラメタの事前分布である。 $p(d)$ は Evidence と呼ばれる量であり、モデル間比較の時には必要になるものの、 θ に依存しないため、パラメタ推定の時には計算の必要がない。

レアイベント探索[‡]

ベイズの定理を日々の天文学研究に応用してみよう。天文学では大量のデータからレアなイベントを探すことが行われる。これを一般にレアイベント探索とよぼう。今、レアイベントを検出する何らかのアルゴリズムを開発しているとする。

レアイベント検出アルゴリズムの性能を

感度：レアイベントを、アルゴリズムがレアイベントを検出したと判定する確率： $p(+|R) = a = 0.5$

特異度：非レアイベント（以降、ゴミと略記）を、アルゴ

リズムがゴミと判定する確率： $p(-|\bar{R}) = b = 0.999$

事象： R =レアイベント、 \bar{R} =ゴミ

事象に対するアルゴリズムの判定： $+$ = アルゴリズムがレアイベントと判定、 $-$ = アルゴリズムがゴミと判定とする。

ここからレアイベント判定されたものが、実際にレアイベントである確率（ここでは簡単に検出確率とよぼう）をレア度 $x = p(R)$ の関数として求めてみよう。

$$\begin{aligned} f(x, a, b) &\equiv p(R|+) = \frac{p(+|R)p(R)}{p(+)} \\ &= \frac{p(+|R)p(R)}{p(+|R)p(R) + p(+|\bar{R})p(\bar{R})} \\ &= \frac{p(+|R)p(R)}{p(+|R)p(R) + [1 - p(-|\bar{R})][1 - p(R)]} \\ &= \frac{ax}{ax + (1-b)(1-x)} \end{aligned} \quad (3.26)$$

レア度が $x = 10^{-4}$ のとき、つまりレアイベントが 10,000 個に 1 個の場合、検出確率は $p(R|+) = f(10^{-4}, 0.5, 0.999) \sim 0.05$ である。この時、まだ 50% しかないアルゴリズムの感度を上げる努力をすべきか、それともすでに 99.9% もあるアルゴリズムの特異度をさらに上げて、ゴミ判定力を強化すべきか？

感度をあげる努力をして、半分しかなかった感度が完璧にレアイベントを検出できるもの、すなわち $a = 0.5$ から $a = 1$ になった場合を考える。このとき検出確率は $p(R|+) = f(10^{-4}, 1, 0.999) \sim 0.09$ となる。元の 2 倍程度、9% の検出確率である。

次にゴミ判定力を向上させ、特異度を $b = 0.999$ から 0.9999 とし、10 倍ゴミを排除できるようにした場合の検出確率は $p(R|+) = f(10^{-4}, 0.5, 0.9999) \sim 0.33$ となり、検出確率は 1/3 くらいにまで上昇する。

これは、イベントが非常にレアで、1 - 特異度（ゴミをレアと誤判定してしまう確率がレア度より大きい場合、つまり $x \ll 1 - b$ の場合）に対応し、この場合、

$$\begin{aligned} f(x, a, b) &= \frac{ax}{ax + (1-b)(1-x)} \\ &\approx \frac{ax}{ax + (1-b)} = \frac{r}{r+1} \sim r \end{aligned} \quad (3.27)$$

$$r \equiv \frac{ax}{1-b} \quad (3.28)$$

となっているので、検出確率を上げるには、通常は既にオーダー 1 となっている感度ではなく、1 - 特異度をレア度（もしくは特異度をゴミ度）に近づける努力のほうが有効である、ということになる。

*2

*2 (つぶやき) レアイベント探索の特性

レアイベント探索では (1 - 特異度)、つまり、ゴミをレアと誤判定してしまう確率を、レア確率 x になるべく近づける努力が重要である。これはゴミについて知れ、という意味であり、モチベーションと相反する。通常、レアイベント探索はレアイベントに興味がある人が行うのであり、ゴミに興味がある人が行うのではない。しかしレアであればあるほど、

3.4 マルコフ鎖モンテカルロ

実際には、尤度関数と事前分布が与えられた場合であっても、解析的にベイズの定理を適用して事後確率を求めることは困難である。天文学データと理論モデルを結び付けるための実際の道具がマルコフ鎖モンテカルロ (MCMC) である。MCMC は、尤度関数・事前分布が与えられた時に、そこから計算される事後確率分布から抽出したサンプリングをモンテカルロ法的に得るアルゴリズムである。

代表的な MCMC のアルゴリズムである Random Metropolis-Hasting (MH) algorithm では、まず初期値として θ_0 をセットした後に、以下の手続きを繰り返すことで、事後確率 $p(\theta|\mathbf{d})$ の定常過程に収束させる。そして事後確率の実現値として、この定常過程から $\{\theta_N, \theta_{N+1}, \dots, \theta_M\}$ がサンプリングされる。

- (1) i -番目の値を θ_i とする。 θ_i から、次のサンプリングの候補 $\hat{\theta}_{i+1}$ (まだ候補なのでハットをつけている) をある確率分布 $q(\hat{\theta}_{i+1}|\theta_i)$ (提案分布) に従うようにランダムに生成する。ここに q はどんな分布でも良い。
- (2) 確率 r で $\hat{\theta}_{i+1}$ を accept する。その場合、 $\theta_{i+1} = \hat{\theta}_{i+1}$ となる。

ここに

$$r(\theta_i, \hat{\theta}_{i+1}) = \min \left[1, \frac{p(\hat{\theta}_{i+1}|\mathbf{d})q(\theta_i|\hat{\theta}_{i+1})}{p(\theta_i|\mathbf{d})q(\hat{\theta}_{i+1}|\theta_i)} \right] \quad (3.29)$$

であり、この値を Metropolis ratio と呼ぶ。このように次の確率が一つ前の結果だけに依存する離散的な

イベント検出の部分は適当でよくなり、ゴミ判定についてよりシビアなアルゴリズムを作らないとならない。

これまでのサーベイで探索していたレア度のものの 100 倍レアなものを見つけるためには、100 倍のサーベイボリュームにするだけでは不十分である。ゴミをレアと誤判定してしまう確率も 1/100 にしなければならず、そのためにはサーベイの精度（ゴミの意味での精度）も 100 倍良くしないとならないかもしれない（元のサーベイのレアイベントがギリギリ見つかった場合）。

また同じレア度のものをサーベイボリュームを 100 倍にしても、1 - 特異度（ゴミをレアと誤判定してしまう確率）が上がってしまうと 100 倍の数は見つからない。データの問題で 1 - 特異度が 100 倍悪くなってしまうと、サーベイボリュームの増加によるゲインはなくなる。

確率過程を Markov chain とよぶため、これらの手法は Markov Chain Monte Carlo (MCMC) と呼ばれる。さて、式 (3.29) の右辺の中を計算するときにベイズの定理を用いる。すなわち

$$r(\theta_i, \hat{\theta}_{i+1}) = \min \left[1, \frac{L(\hat{\theta}_{i+1})p(\hat{\mathbf{p}}_{i+1})}{L(\theta_i)p(\theta_i)} \right] \quad (3.30)$$

のように尤度と事前分布を与えれば $r(\theta_i, \hat{\theta}_{i+1})$ を計算できることがわかる。ただしここで提案分布は対称、すなわち $q(\theta_i|\hat{\theta}_{i+1}) = q(\hat{\theta}_{i+1}|\theta_i)$ となるもの、例えばガウシアンなど、を選んだとしている。

さて上の手続きで定常分布が成り立ち、かつそれが $p(\theta|\mathbf{d})$ となることを示そう。 i 番目で $p(\theta_i|\mathbf{d})$ に従っているとして、その確率分布が、手続きにより確率 $p(\theta_{i+1}|\theta_i)$ で θ_{i+1} を生成するとき、 $i+1$ 番目で θ_{i+1} は

$$p(\theta_{i+1}) = \int p(\theta_{i+1}|\theta_i)p(\theta_i|\mathbf{d})d\theta_i \quad (3.31)$$

の確率に従うであろう。定常分布であるためにはこの $p(\theta_{i+1})$ が再び $p(\theta_{i+1}|\mathbf{d})$ と一致すれば良い。このために必要な条件が詳細釣り合い条件

$$p(\theta_{i+1}|\theta_i)p(\theta_i|\mathbf{d}) = p(\theta_i|\theta_{i+1})p(\theta_{i+1}|\mathbf{d}) \quad (3.32)$$

である。詳細釣り合いが成り立てば、

$$\begin{aligned} p(\theta_{i+1}) &= \int p(\theta_{i+1}|\theta_i)p(\theta_i|\mathbf{d})d\theta_i \\ &= \int p(\theta_i|\theta_{i+1})p(\theta_{i+1}|\mathbf{d})d\theta_i = p(\theta_{i+1}|\mathbf{d}) \end{aligned} \quad (3.33)$$

となる。

さて、MH algorithm では θ_i から θ_{i+1} が生成される確率は

$$p(\theta_{i+1}|\theta_i) = r(\theta_i, \theta_{i+1}) q(\theta_{i+1}|\theta_i) \quad (3.34)$$

であるため、

$$\begin{aligned}
p(\boldsymbol{\theta}_{i+1}|\boldsymbol{\theta}_i)p(\boldsymbol{\theta}_i|\mathbf{d}) &= r q(\boldsymbol{\theta}_{i+1}|\boldsymbol{\theta}_i)p(\boldsymbol{\theta}_i|\mathbf{d}) \\
&= \min \left[1, \frac{p(\boldsymbol{\theta}_{i+1}|\mathbf{d})q(\boldsymbol{\theta}_i|\boldsymbol{\theta}_{i+1})}{p(\boldsymbol{\theta}_i|\mathbf{d})q(\boldsymbol{\theta}_{i+1}|\boldsymbol{\theta}_i)} \right] q(\boldsymbol{\theta}_{i+1}|\boldsymbol{\theta}_i)p(\boldsymbol{\theta}_i|\mathbf{d}) \\
&= \min [q(\boldsymbol{\theta}_{i+1}|\boldsymbol{\theta}_i)p(\boldsymbol{\theta}_i|\mathbf{d}), p(\boldsymbol{\theta}_{i+1}|\mathbf{d})q(\boldsymbol{\theta}_i|\boldsymbol{\theta}_{i+1})] \\
&= \min \left[\frac{p(\boldsymbol{\theta}_i|\mathbf{d})q(\boldsymbol{\theta}_{i+1}|\boldsymbol{\theta}_i)}{p(\boldsymbol{\theta}_{i+1}|\mathbf{d})q(\boldsymbol{\theta}_i|\boldsymbol{\theta}_{i+1})}, 1 \right] q(\boldsymbol{\theta}_i|\boldsymbol{\theta}_{i+1})p(\boldsymbol{\theta}_{i+1}|\mathbf{d}) \\
&= r(\boldsymbol{\theta}_{i+1}, \boldsymbol{\theta}_i)q(\boldsymbol{\theta}_i|\boldsymbol{\theta}_{i+1})p(\boldsymbol{\theta}_{i+1}|\mathbf{d}) \\
&= p(\boldsymbol{\theta}_i|\boldsymbol{\theta}_{i+1})p(\boldsymbol{\theta}_{i+1}|\mathbf{d}) \tag{3.35}
\end{aligned}$$

となり、たしかに詳細釣り合いが成り立っていることが確かめられる。ただし、初期条件として適当な $\boldsymbol{\theta}_0$ から始めるが、Markov chain の最初の部分は初期条件に依存しているので解析から除くことが必要である。

3.5 ハミルトニアン・モンテカルロと自動微分

random MH の問題点は提案分布によるランダムな移動を伴うため、高次元で棄却率が高くなってしまいう点である。ハミルトニアンモンテカルロ (HMC) は、求めたいパラメタ $\boldsymbol{\theta}$ に対し、対応する“運動量” \mathbf{p} を導入し、ハミルトニアン保存を利用して棄却率を下げる手法である。ここでポテンシャルエネルギー U と運動エネルギー K を

$$U(\boldsymbol{\theta}) = -\log p(\boldsymbol{\theta}|\mathbf{d}) \tag{3.36}$$

$$K(\mathbf{p}) = \frac{1}{2}\mathbf{p}^\top M\mathbf{p} \tag{3.37}$$

とする。ここに M は mass matrix とよばれる正定値行列である。ハミルトニアンを

$$H(\boldsymbol{\theta}, \mathbf{p}) \equiv U(\boldsymbol{\theta}) + K(\mathbf{p}) \tag{3.38}$$

と定義し、同時確率分布を

$$p(\boldsymbol{\theta}, \mathbf{p}|\mathbf{d}) = e^{-H(\boldsymbol{\theta}, \mathbf{p})} = p(\boldsymbol{\theta}|\mathbf{d})p(\mathbf{p}) \tag{3.39}$$

$$p(\mathbf{p}) = \exp\left(-\frac{1}{2}\mathbf{p}^\top M\mathbf{p}\right) \tag{3.40}$$

とする。つまり mass matrix はガウシアン共分散の逆数 $M = \Sigma^{-1}$ であることになる。 $(\boldsymbol{\theta}, \mathbf{p})$ がハミルトン方程式に従うとすると、ハミルトニアンは時間保

存量となる。すなわち

$$\dot{\boldsymbol{\theta}} = \frac{\partial H(\boldsymbol{\theta}, \mathbf{p})}{\partial \mathbf{p}} = \frac{\partial K(\mathbf{p})}{\partial \mathbf{p}} \tag{3.41}$$

$$\dot{\mathbf{p}} = -\frac{\partial H(\boldsymbol{\theta}, \mathbf{p})}{\partial \boldsymbol{\theta}} = -\frac{\partial U(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \tag{3.42}$$

ならば、

$$\dot{H} = \sum_{i=1}^D \left[\frac{\partial H(\boldsymbol{\theta}, \mathbf{p})}{\partial p_i} \dot{p}_i + \frac{\partial H(\boldsymbol{\theta}, \mathbf{p})}{\partial q_i} \dot{q}_i \right] \tag{3.43}$$

$$= \frac{\partial H(\boldsymbol{\theta}, \mathbf{p})}{\partial \boldsymbol{\theta}} \dot{\boldsymbol{\theta}} + \frac{\partial H(\boldsymbol{\theta}, \mathbf{p})}{\partial \mathbf{p}} \dot{\mathbf{p}} = 0 \tag{3.44}$$

となる。

さてここでメトロポリス・ヘイスティングスの受容率 (3.29) は

$$\begin{aligned}
r(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}_{i+1}) &= \min \left[1, \frac{p(\hat{\boldsymbol{\theta}}_{i+1}, \hat{\mathbf{p}}_{i+1}|\mathbf{d})}{p(\boldsymbol{\theta}_i, \mathbf{p}_i|\mathbf{d})} \right] \\
&= \min \left[1, e^{H(\boldsymbol{\theta}_i, \mathbf{p}_i) - H(\hat{\boldsymbol{\theta}}_{i+1}, \hat{\mathbf{p}}_{i+1})} \right] \tag{3.45}
\end{aligned}$$

となる。つまり力学計算を精度良く行い、ハミルトニアンがほぼ保存すれば、受容率はほぼ 1 となる。最終的に $p(\boldsymbol{\theta}, \mathbf{p}|\mathbf{d})$ のサンプリングが行われれば、 \mathbf{p} について周辺化すれば $p(\boldsymbol{\theta}|\mathbf{d})$ のサンプリングが得られることに注意する。

詳しく解説はしないが、力学計算は Leap-frog を用いて行う。この際、 $U(\boldsymbol{\theta})$ の $\boldsymbol{\theta}$ による勾配が必要である。これは尤度 $p(\mathbf{d}|\boldsymbol{\theta})$ の $\boldsymbol{\theta}$ 微分が必要である。尤度関数の中にモデルが含まれることから、最終的にモデルの $\boldsymbol{\theta}$ が必要であることがわかる。すなわち、HMC にはモデルの勾配計算が必要である。

数値微分は誤差が蓄積するため、通常、不適である。また手計算で微分を与えてもよいが、複雑なモデルの場合、フレキシビリティの観点からも自動微分が用いられることが多い。つまり HMC を行うためには微分可能プログラミングでモデルを構築することが肝要である。

3.6 ガウス過程[‡]

N 次元ガウシアン分布を

$$\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \Sigma) \equiv \frac{1}{\sqrt{(2\pi)^N |\Sigma|}} e^{-(\mathbf{x}-\boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{x}-\boldsymbol{\mu})/2} \tag{3.47}$$

と表記する。また誤解がない場合、変数 \mathbf{x} を略して、

$$\mathcal{N}(\boldsymbol{\mu}, \Sigma) \equiv \frac{1}{\sqrt{(2\pi)^N |\Sigma|}} e^{-(\mathbf{x}-\boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{x}-\boldsymbol{\mu})/2} \quad (3.48)$$

のように表記する場合もある。

相関ノイズとしてのガウス過程

ガウス過程は、多変数正規分布

$$\mathbf{d} \sim \mathcal{N}(\mathbf{0}, \Sigma) \quad (3.49)$$

に従う確率変数 \mathbf{x} の確率過程のことである。いま d_i は時系列だとして、この多変数正規分布の共分散行列の各成分が、

$$\Sigma_{ij} = K_{ij}(a, \tau) = ak(|t_i - t_j|; \tau) \quad (3.50)$$

のように、 t_i と t_j の差の絶対値の関数として与えられれば、相関長 τ のガウス過程が得られる。カーネル関数 $k(t, \tau)$ としてはいろいろなタイプがありうるが、例えば RBF カーネル

$$k_{\text{RBF}}(t; \tau) = \exp\left(-\frac{t^2}{2\tau^2}\right), \quad (3.51)$$

と Matérn 3/2 カーネル

$$k_{\text{M3/2}}(t; \tau) = \left(1 + \frac{\sqrt{3}t}{\tau}\right) e^{-\sqrt{3}t/\tau}. \quad (3.52)$$

などがある。

このガウス過程にしたがったデータ点をサンプリングしてみると図3.2のようになる。ただしデータにさらに平均ゼロ、標準偏差 σ のガウスノイズを足している。このノイズを便宜的に観測ノイズと呼んでおこう。

さてこのデータを元に τ 、 a 、 σ を MCMC サンプリングしてみよう。平均ゼロ、標準偏差 σ のガウスノイズの観測ノイズを足したということは確率モデルとしては、

$$\mathbf{d} \sim \mathcal{N}(\mathbf{0}, \Sigma') \quad (3.53)$$

$$\Sigma' = K(a, \tau) + \sigma^2 I \quad (3.54)$$

となる。この Σ' は RBF カーネルを仮定し、さらに独立誤差を表す対角項を加える。つまり

$$\Sigma_{ij}(a, \tau, \text{sigma}) = ak_{\text{RBF}}(|t_i - t_j|; \tau) + \sigma^2 I \quad (3.55)$$

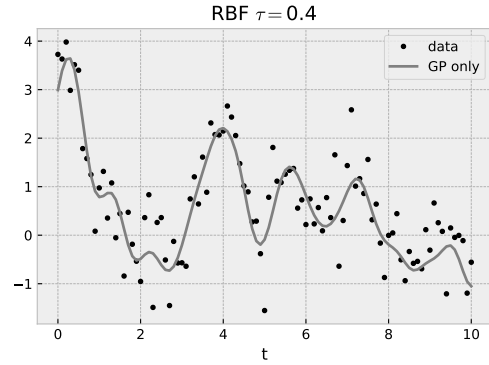


図3.2 ガウス過程からサンプルされたデータ。実線は σ の観測ノイズなし。

となり、共分散はパラメタ (a, τ, σ) を持つ。Exponential 分布 $E(x)$ を用いて事前分布を構成すると、たとえば事前分布は

$$p(a) = E(1) \quad (3.56)$$

$$p(\tau) = E(1) \quad (3.57)$$

$$p(\sigma) = E(1) \quad (3.58)$$

となり、尤度は

$$p(\mathbf{d}|a, \tau, \sigma) = \mathcal{N}(\mathbf{0}, \Sigma(a, \tau, \sigma)) \quad (3.59)$$

のようにかける。つまりこれで MCMC を回すことができる。

HMC でサンプルした事後分布を可視化したものが図3.3である。ところで、式 (3.53) という観点のモデル化では、 $\mathbf{x} = \mathbf{0}$ がモデル平均である。credible interval を計算したものが図3.4となっているが、このことが直感的にわかる。

モデルとしてのガウス過程

上記のガウス過程による解析は、モデルパラメタ \mathbf{m} の事前分布（プライア）として

$$p(\mathbf{m}) = \mathcal{N}(\mathbf{0}, \Sigma) \quad (3.60)$$

と置き、

$$d_i = m_i + \epsilon \quad (3.61)$$

$$\epsilon \sim \mathcal{N}(0, \sigma^2) \quad (3.62)$$

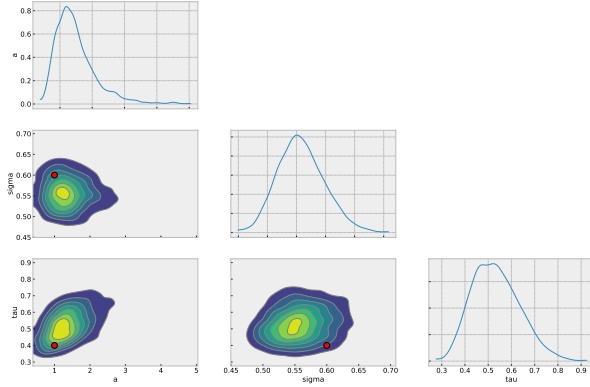


図3.3

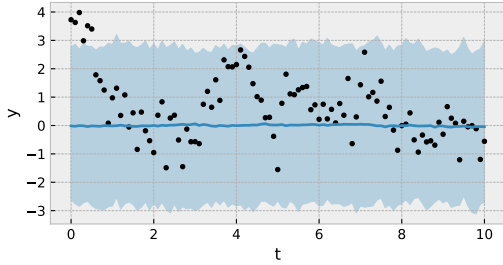


図3.4

と見ることもできる。ここに ϵ は観測ノイズに対応している。もう少しまどろっこしく書くとモデル \mathbf{g} を恒等変換として、

$$\mathbf{g}(\mathbf{m}) = \mathbf{m} \quad (3.63)$$

$$\mathbf{d} = \mathbf{g}(\mathbf{m}) + \epsilon \quad (3.64)$$

$$\epsilon \sim \mathcal{N}(\mathbf{0}, \sigma^2 I) \quad (3.65)$$

のように書ける。この場合、尤度関数が

$$p(\mathbf{d}|\mathbf{m}) = \mathcal{N}(\mathbf{d}; \mathbf{g}(\mathbf{m}), \sigma^2 I) = \mathcal{N}(\mathbf{d}; \mathbf{m}, \sigma^2 I) \quad (3.66)$$

となる。そして、さらにこの事前分布の中にパラメタがあるという構造となっている。つまり

$$\Sigma_{ij} = ak(|t_i - t_j|; \tau) \quad (3.67)$$

としていて、このパラメタ a, τ をハイパーパラメタという。そしてハイパーパラメタの事前分布 (超事前分

布またはハイパープライア) をさらに仮定するという事に相当する。

ハイパーパラメタを固定した状態では、 \mathbf{m} の事後分布を解析的に書けることが知られている。ここでガウス過程の計算法を紹介しよう。まず多変数正規分布に多変数正規分布をかけても多変数正規分布である。そこで計算しなくてはならないのは \exp のべきの部分のみである。多変数正規分布の「べき」部分は

$$\begin{aligned} -2 \log \mathcal{N}(\mathbf{m}; \boldsymbol{\mu}, \Sigma) &= (\mathbf{m} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{m} - \boldsymbol{\mu}) \\ &= \mathbf{m}^\top \Sigma^{-1} \mathbf{m} - 2\mathbf{m}^\top \Sigma^{-1} \boldsymbol{\mu} + \text{const.} \end{aligned} \quad (3.68)$$

となっていることに注意すると、もしある多変数正規分布に従うことがわかっている確率密度分布 $p(\mathbf{m})$ が

$$-2 \log p(\mathbf{m}) = \mathbf{m}^\top P \mathbf{m} - 2\mathbf{m}^\top \mathbf{q} + \text{const.}, \quad (3.69)$$

のように書けたとすると、式 (3.68) と比べることで、この確率密度分布は

$$p(\mathbf{m}) = \mathcal{N}(\mathbf{m}; P^{-1} \mathbf{q}, P^{-1}). \quad (3.70)$$

であることがわかる。

さて、いま事後分布は

$$p(\mathbf{m}|\mathbf{d}) \propto p(\mathbf{d}|\mathbf{m})p(\mathbf{m}) = \mathcal{N}(\mathbf{d}; \mathbf{m}, \sigma^2 I) \mathcal{N}(\mathbf{m}; \mathbf{0}, \Sigma) \quad (3.71)$$

であるので、この「べき」部分を \mathbf{m} について展開することで、

$$\begin{aligned} -2 \log [\mathcal{N}(\mathbf{d}; \mathbf{m}, \sigma^2 I) \mathcal{N}(\mathbf{m}; \mathbf{0}, \Sigma)] \\ = \mathbf{m}^\top (\Sigma^{-1} + \sigma^{-2} I) \mathbf{m} - 2\sigma^{-2} \mathbf{m}^\top \mathbf{d} + \text{const.} \end{aligned} \quad (3.72)$$

と書けること、また

$$(\Sigma^{-1} + \sigma^{-2} I)^{-1} = \Sigma(I + \sigma^{-2} \Sigma)^{-1} \quad (3.73)$$

より

$$p(\mathbf{m}|\mathbf{d}) = \mathcal{N}(\mathbf{m}; \Sigma(\sigma^2 I + \Sigma)^{-1} \mathbf{d}, \Sigma(I + \sigma^{-2} \Sigma)^{-1}) \quad (3.74)$$

であることがわかった。

MCMC により τ と σ のサンプリングをすることができるが、これらのパラメタの事後分布は何を意味

するか考えよう。これらは今の枠組みではハイパーパラメタとみなせるので、まとめて $\theta = (a, \tau, \sigma)^\top$ をハイパーパラメタとおく。ここからは θ も確率に入れて考えることにする。さて、今、 \mathbf{m} について周辺化した尤度

$$p(\mathbf{d}|\theta) = \frac{p(\mathbf{d}|\mathbf{m}, \theta)p(\mathbf{m}, \theta)}{p(\mathbf{m}|\mathbf{d}, \theta)} \quad (3.75)$$

はやはり多変数正規分布となるが、また「べき」部分の \mathbf{d} に関する部分だけ考えることにより

$$\begin{aligned} -2 \log p(\mathbf{d}|\theta) &= \\ -2 \log p(\mathbf{d}|\mathbf{m}, \theta) + 2 \log p(\mathbf{m}|\mathbf{d}, \theta) + \text{const} \\ &= -2 \log \mathcal{N}(\mathbf{d}|\mathbf{m}, \sigma^2 I) \\ &\quad + 2 \log \mathcal{N}(\mathbf{m}|\Sigma(\sigma^2 I + \Sigma)^{-1}\mathbf{d}, \Sigma(I + \sigma^{-2}\Sigma)^{-1}) \\ &= \mathbf{d}^\top (\sigma^2 I + \Sigma)^{-1} \mathbf{d} + \text{const.} \end{aligned} \quad (3.76)$$

である。よって

$$p(\mathbf{d}|\theta) = \mathcal{N}(\mathbf{d}; \mathbf{0}, \Sigma + \sigma^2 I) \quad (3.77)$$

となる。 $\Sigma = K(\tau)$ とすれば、式 (3.53) と同じである。すなわち、(超)事前分布として $p(\theta)$ を与えた時の周辺事後分布

$$p(\theta|\mathbf{d}) \propto p(\mathbf{d}|\theta)p(\theta) \quad (3.78)$$

をサンプリングしていることになる。

図3.3はハイパーパラメタ τ, σ の周辺事後分布を示していることがわかった。つまり $k = 0, \dots, N_s - 1$ に対して

$$\theta_k^\dagger \sim p(\theta|\mathbf{d}) \propto p(\mathbf{d}|\theta)p(\theta) \quad (3.79)$$

をサンプリングしたことになる。さて、ではこのサンプルと式 (3.74) を用いて、

$$p(\mathbf{m}, \theta|\mathbf{d}) = p(\mathbf{m}|\theta, \mathbf{d})p(\theta|\mathbf{d}) \quad (3.80)$$

であることから

$$p(\mathbf{m}|\theta_k^\dagger, \mathbf{d}) = \mathcal{N}(\mu_k, K_k) \quad (3.81)$$

$$\mu_k = K(a_k^\dagger, \tau_k^\dagger)((\sigma_k^\dagger)^2 I + K(a_k^\dagger, \tau_k^\dagger))^{-1} \mathbf{d} \quad (3.82)$$

$$K_k = K(a_k^\dagger, \tau_k^\dagger)(I + (\sigma_k^\dagger)^{-2} K(a_k^\dagger, \tau_k^\dagger))^{-1} \quad (3.83)$$

からサンプルした \mathbf{m}_k^\dagger と θ_k^\dagger のセットは

$$(\mathbf{m}_k^\dagger, \theta_k^\dagger) \sim p(\mathbf{m}, \theta|\mathbf{d}) \quad (3.84)$$

とみなせることがわかる*3。

HPDI を計算することで、図3.5の \mathbf{m} の credible interval が求まる。ここで注意点としては、これはモデル \mathbf{m} の 90% interval であり、観測ノイズ部分を含んだ \mathbf{d} の予測ではないという点である。

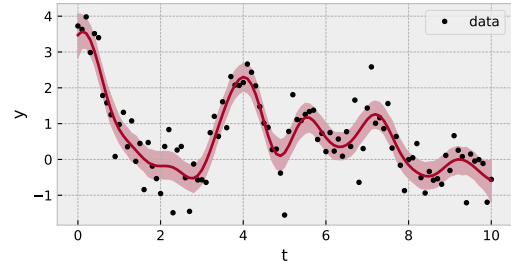


図3.5

データ点にない位置の予測

さて、一般に $t = t^*$ での予測値はどうなるだろうか？以下では再度ハイパーパラメタを省略して表記する。まず、 \mathbf{m}, \mathbf{m}^* がガウス過程に従うとき

$$p(\mathbf{m}^*|\mathbf{m}) = \mathcal{N}(K_\times^\top K^{-1} \mathbf{m}, K_* - K_\times^\top K^{-1} K_\times) \quad (3.85)$$

である。ここに

$$K_{ij} = ak(|t_i - t_j|; \tau) \quad (3.86)$$

$$(K_\times)_{ij} = ak(|t_i - t_j^*|; \tau) \quad (3.87)$$

$$(K_*)_{ij} = ak(|t_i^* - t_j^*|; \tau) \quad (3.88)$$

同様に

$$p(\mathbf{m}^*|\mathbf{d}) = \mathcal{N}(K_\times^\top K_\sigma^{-1} \mathbf{d}, K_* - K_\times^\top K_\sigma^{-1} K_\times) \quad (3.89)$$

*3 ハイパーパラメタの点推定 (いわゆる maximum marginal likelihood=maximum evidence) を用いる方法については解説が多くあるが、周辺事後分布から再サンプリングする議論についてはいまのところ文献が見つからない。ベイズ線形問題+ガウス過程の枠組みで以前、筆者が議論したことがあるので参考までに挙げておく [6]。

となる。ここに

$$(K_\sigma)_{ij} = ak(|t_i - t_j|; \tau) + \sigma^2 \delta_{i,j} \quad (3.90)$$

($\delta_{i,j}$ はクロネッカーデルタ)。ここで $p(\mathbf{m}^*|\mathbf{d})$ はあくまでモデルパラメタとしてのガウス過程の事後分布なので、観測ノイズの分は考慮されていない。

もし観測ノイズを含んだ予測を行いたいのであれば、

$$\mathbf{d}^* = \mathbf{m}^* + \epsilon \quad (3.91)$$

というモデルに基づき、

$$p(\mathbf{d}^*|\mathbf{d}) = \mathcal{N}(K_\times^\top K_\sigma^{-1} \mathbf{d}, K_{*,\sigma} - K_\times^\top K_\sigma^{-1} K_\times) \quad (3.92)$$

となるだろう。ここに

$$(K_{*,\sigma})_{ij} = ak(|t_i^* - t_j^*|; \tau) + \sigma^2 \delta_{i,j} \quad (3.93)$$

である。

というわけで、ハイパーパラメタ込みのサンプリングで式 (3.89, 3.92) のサンプリングを行い Credible Interval を求める。同様に HPDI を求めたものが図3.6である。濃い色が \mathbf{m}^* の、薄い色が \mathbf{d}^* の credible interval である。

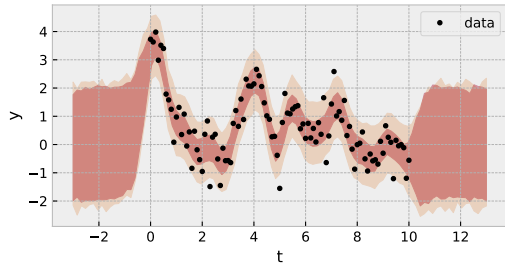


図3.6

3.7 ガウス過程ノイズを含んだモデルフィット ‡

さて、ここまではモデルとしては平均値がゼロのガウス過程により生成されるモデル

$$\mathbf{m} \sim \mathcal{N}(\mathbf{0}, \Sigma(t)) \quad (3.94)$$

を考えていた。ここでは一般に平均値が t の関数となるガウス過程によりデータが生成される場合を考えよう。すなわち

$$\mathbf{m} \sim \mathcal{N}(\mathbf{f}(t), \Sigma(t)) \quad (3.95)$$

の場合を考えよう。ここに $\mathbf{f}(t)$ は $f(t)$ に対し、 $\mathbf{t} = (t_0, t_1, \dots, t_{N-1})$ を要素ごとに適用したベクトル版である。ここでは $f(t)$ として

$$f(t) = ke^{-(t-T_0)^2/2s^2} \sin(2\pi t/P) \quad (3.96)$$

というものを考えてみよう。上では簡単に $f(t)$ と言いたが、パラメタを $\theta = (T_0, k, s, P)$ として $f(t; \theta)$ という表記も併用する。

式 (3.95) の共分散行列を RBF カーネルとして選択すると、例えば、図3.7上のようなデータが生成される。ここで $\tau = 3$ であり $f(t)$ より比較的ゆったりとしたトレンドが乗っているのが見て取れるだろう。このようなモデルを HMC フィットするというのは、

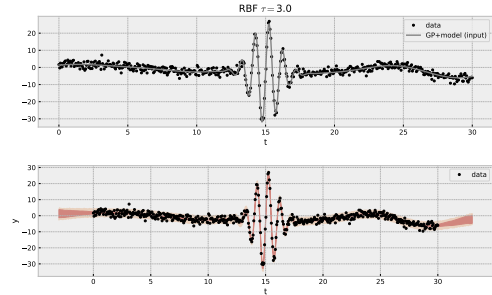


図3.7 実線は σ の観測ノイズなし。

シグナル $f(t)$ と相関のあるノイズ+観測ノイズをモデル化し、 $f(t)$ の持つパラメタを推定したいときに対応する。

$\mathbf{f}(t)$ がわかっている場合、観測ノイズを含まない、および、含んだ予測は、それぞれ、

$$p(\mathbf{m}^*|\mathbf{d}) = \mathcal{N}(\mathbf{f}(t^*) + K_\times^\top K_\sigma^{-1}(\mathbf{d} - \mathbf{f}(t)), K_* - K_\times^\top K_\sigma^{-1} K_\times) \quad (3.97)$$

$$p(\mathbf{d}^*|\mathbf{d}) = \mathcal{N}(\mathbf{f}(t^*) + K_\times^\top K_\sigma^{-1}(\mathbf{d} - \mathbf{f}(t)), K_{*,\sigma} - K_\times^\top K_\sigma^{-1} K_\times) \quad (3.98)$$

となる。 $\mathbf{f}(t)$ の持つパラメタ $\boldsymbol{\theta} = (T_0, k, s, P)$ は HMC でサンプリングされるから、例えば後者なら、サンプリングされた各 $\boldsymbol{\theta}_k^\dagger$ をもちいて

$$\mathbf{d}_k^* \sim \mathcal{N}(\mu_k, K_k) \quad (3.99)$$

$$\mu_k = \mathbf{f}(\mathbf{t}^*; \boldsymbol{\theta}_k^\dagger) + K_\times^\top K_\sigma^{-1}(\mathbf{d} - \mathbf{f}(\mathbf{t}; \boldsymbol{\theta}_k^\dagger)) \quad (3.100)$$

$$K_k = K_{*,\sigma} - K_\times^\top K_\sigma^{-1} K_\times \quad (3.101)$$

をサンプリングすれば、予測のサンプリングができる。というわけで、以下のようにモデル $\mathbf{f}(t)$ 込みの GP のフィットを行うことができる (図3.7下)。

第 4 章

系外惑星からのスペクトル

本章では惑星自身のもつ情報が埋め込まれたシグナルの概観を行う。惑星シグナルは、大まかに分けると恒星光が惑星大気を透過したもの、惑星からの熱輻射、もしくは恒星光が惑星により反射されたもののいずれかの形態をとる (表4.1)。これらの惑星シグナルを分光することで系外惑星の大気を調べることができる。いずれの場合も、系外惑星大気を通過する光が、大気中の分子・原子や雲といった成分によって波長方向に吸収や散乱といった影響を受けて変化する。その変化をスペクトルを通じて検出することで大気を理解するということである。通過する光は上の三つの形態に応じて惑星自身が放つ輻射光と恒星光の光を利用する方法に分けることができる。前者は惑星からの輻射光という意味で今後「輻射光」と略記する。後者はそれぞれトランジット透過光と反射光であり、この二者は恒星光に惑星シグナルが刻まれるという意味でパッシブな手法であるとも見ることができる。

表4.1 スペクトル種類の分類

スペクトル	手法	現在の主な対象
透過光	透過光分光	トランジット系
	高分散分光	トランジット系
輻射	二次食分光	トランジット系
	直接撮像	若い惑星
	高分散分光	明るい系
反射	直接撮像	地球近傍
	高分散分光	明るい系

系外惑星分光の最大の問題は、いかに恒星光の影響を分離・除去し、惑星光のスペクトルを取り出すかというところにある。表4.2に、分離方法に基づいた手法の分類を示す。まずデータ解析で分離すると手法は、恒星光と惑星光がまったく分離されていないスペ

クトルを取得し、データ解析的に惑星光を分離する手法である。トランジット系では、トランジット時の減光から、惑星の半径が計測できるが、これを波長ごとに行い、大気情報を得ることができる。この手法は透過光分光と呼ばれる。トランジット系では、惑星が恒星の背面に隠れる時に、惑星光が隠された分やはり減光する。これは二次食と呼ばれ、惑星光の分離に用いることができる。この二つの手法はトランジット系のみに行うことができる方法である。次に恒星光と惑星光が混ざったものに対し高分散分光を利用し惑星光を取り出す手法がある。この手法では恒星と惑星の視線速度差を利用して、惑星由来の成分、具体的には、惑星大気の子分子や原子の吸収線を同定する。この手法は非トランジット系にも適用可能である。最後に、装置的に恒星光を除去する手法が直接撮像法である。この手法は、主に恒星光と惑星光の光の強度比(コントラスト)と恒星と惑星の分離角によって検出できるかどうかが決まる。分離角は観測者に近いほど相対的に大きくなるので、地球近傍の惑星系がターゲットとなる。また現在の直接撮像の性能では、コントラストが穏やかなものしか検出できないため、惑星が高温で自ら光っている若い惑星系のみが検出されている。ただし、今後の宇宙直接撮像や地上の装置性能の向上に伴い、原理的には通常の惑星系でも検出可能である。

表4.2 系外惑星分光法の分類

分離手段	手法	対称
データ解析	透過光分光	トランジット惑星
	二次食分光	トランジット惑星
	高分散分光	トランジット or 高温惑星
装置	直接撮像	若い惑星 or 地球近傍

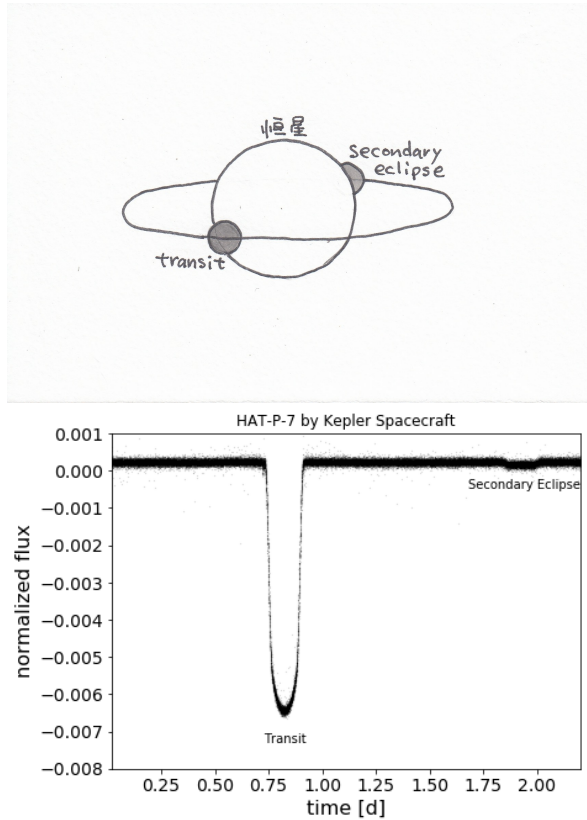


図4.1 トランジット系の概念図（上）とホットジュピター HAT-P-7b によるトランジット減光曲線の例（下）。

4.1 トランジット現象

主星フラックスの変動にも惑星によるシグナルが含まれる。最も分かりやすいのは主星の前面を惑星が通過する時にできる陰によるフラックス減少であるトランジット減光である（図4.1）。トランジット系は逆に惑星が主星の後ろに隠れることによる減光も観測されている。後者は二次食 (secondary eclipse) とか単に occlusion と呼ばれている。

惑星がトランジットする確率は、円軌道かつ $R_p \ll R_*$ を仮定すると、図4.2で示された領域であるから

$$p_{\text{tra}} = \sin(R_*/a) \sim \frac{R_*}{a} = 0.005 \left(\frac{R_*}{R_\odot} \right) \left(\frac{a}{1 \text{ au}} \right)^{-1} \quad (4.1)$$

となり、G 型星まわりのハビタブル惑星だと 0.5 %

である。

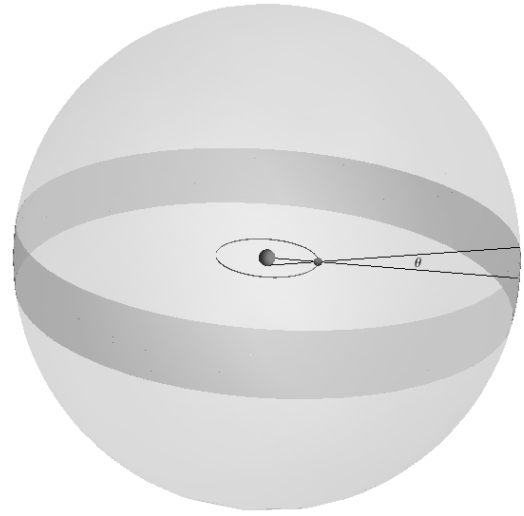


図4.2 トランジットして見える立体角方向（帯部分）。中に恒星と惑星軌道が描かれている。外側の球の半径は、観測者から惑星系までの距離 d であるので、 $a \ll d$ と見なせる。するとランダムな視線方向で帯領域をとる確率は $4\pi \sin \theta / (4\pi) \approx R_*/a$ となる。

トランジットライトカーブ

さてトランジットのライトカーブから、系外惑星系のこういった物理量が分かるのだろうか？まず複数回のトランジットが観測されれば、

- 公転周期 P

を知ることができる。また、トランジット深さは、

- 惑星半径と恒星半径の比 $k \equiv R_p/R_*$

の二乗となるので、この k を知ることができる。トランジットの継続時間から

- 惑星の通過時間 (duration) T_{tot}

を知ることができる。さらに細かく見ると、もし恒星が一様に光っているという近似の元では、図4.3a に示されるように、ingress・egress 時もわかる。つまり惑星が恒星に入るとき出る時に、惑星が恒星に入り始めてから、完全に出るまでの時間 t_T と、惑星が完全に

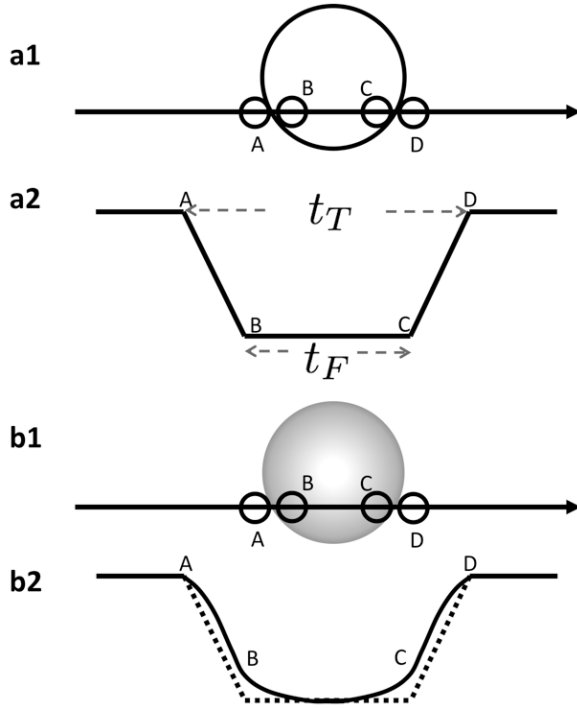


図4.3 トランジットライトカーブの幾何。a は恒星が一樣に光っている場合のライトカーブ。実際は恒星は淵側が暗い (周辺減光) ために b のようなライトカーブとなる。

恒星の全面に入っている期間、 t_F が測定できるというように言い換えることもできよう。恒星は、実際は周辺減光 (Limb darkening) をしており、実際のライトカーブは図4.3b のようになる。そこで通常、周辺減光のモデル化を行うことで t_T と t_F も求めることができる。周辺減光のモデルとしては以下の quadratic limb-darkening law

$$I(\mu) = I(\mu = 1)[1 - u_1(1 - \mu) - u_2(1 - \mu)^2] \quad (4.2)$$

がよく用いられる。ここに $\mu = \cos \psi$ であり ψ は光球面の法線ベクトルと視線ベクトルのなす角度である。つまり中心で $\mu = 1$ 、ディスク境界で $\mu = 0$ となる。上記モデルは円盤で積分すると $\pi R_*^2(1 - u_1/3 - u_2/6)I(\mu = 1)$ になることに注意。今、 k がわかっている、 t_T , t_F から幾何学的に

- 衝突径数 (impact parameter): $b = \frac{a}{R_*} \cos i$

も知ることができる。恒星スペクトルのモデルや後述する恒星密度 ρ_* の情報を用いて R_* と M_* が推定できれば、周期 P がわかっているとケプラー第三法則から a を推定することができる。これにより軌道傾斜角 i を決定することができる。もし視線速度の測定されているトランジット系が存在すれば、軌道傾斜角 i がわかっているので惑星質量 M_p が求まるので、惑星の平均密度

$$\rho_p \equiv \frac{3M_p}{4\pi R_p^3} \quad (4.3)$$

が推定できるということである。これにより系外惑星の大まかな分類が可能になった。

円軌道の場合、通過時間は $T_{\text{tot}} = 2\sqrt{1 - b^2}R_*/v$ である。ここに v は惑星の通過速度である。惑星の質量を無視したケプラー第三法則

$$P^2 = \frac{4\pi^2}{GM_*}a^3 \quad (4.4)$$

を $v = 2\pi a/P$ であることに注意して変形すると

$$v^3 = 2\pi GM_*/P \quad (4.5)$$

となるので、 $T_{\text{tot}}^3 \propto P/\rho_*$ となり、周期・衝突径数がわかっていると、通過時間からは恒星密度 ρ_* がわかることになる。逆にだいたいの通過時間を知っておくため $b = 0$ のときを書いてみよう。この場合、

$$T_{\text{tot}} = \left(\frac{3P}{\pi^2 G \rho_*} \right)^{1/3} \quad (4.6)$$

$$= 2.6\text{h} \left(\frac{P}{3\text{day}} \right)^{1/3} \left(\frac{\rho_*}{\rho_\odot} \right)^{-1/3} \quad (4.7)$$

$$= 30\text{h} \left(\frac{P}{12\text{yr}} \right)^{1/3} \left(\frac{\rho_*}{\rho_\odot} \right)^{-1/3} \quad (4.8)$$

となる。二行目はホット・ジュピターの典型値で三行目は木星の場合を仮定している。このように周期が数年以上になっていくと、通過時間が一日を超えるようになってくる。また、巨星の場合は、恒星密度が桁で小さくなるので通過時間がやはり長くなる。

column – 周期とライトカーブ形状 †

円軌道を仮定すると幾何的考察から、

$$\sin\left(\frac{\pi t_T}{P}\right) = \frac{R_*}{a} \sqrt{\frac{(1+k)^2 - b^2}{\sin^2 i}}, \quad (4.9)$$

$$\sin\left(\frac{\pi t_F}{P}\right) = \frac{R_*}{a} \sqrt{\frac{(1-k)^2 - b^2}{\sin^2 i}}, \quad (4.10)$$

が得られる。ここで、 $\sin(\pi t_T/P) \sim \pi t_T/P$ 、 $\sin(\pi t_F/P) \sim \pi t_F/P$ と近似すれば、

$$\frac{R_*}{a} = \frac{\pi}{2\sqrt{k}} \frac{\sqrt{t_T^2 - t_F^2}}{P} \sin i. \quad (4.11)$$

が得られる。さらに惑星の質量を無視したケプラー第三法則

$$P^2 = \frac{4\pi^2}{GM_*} a^3 \quad (4.12)$$

を用い $\sin i \sim 1$ とすると、

$$P = \frac{\pi G M_*}{32 R_*^3} \left(\frac{t_T^2 - t_F^2}{k} \right)^{\frac{3}{2}} \quad (4.13)$$

$$= \frac{\pi^2 G}{24} \rho_* \left(\frac{t_T^2 - t_F^2}{k} \right)^{\frac{3}{2}} \quad (4.14)$$

が得られる。最後の式で未知の量は、恒星の平均密度 ρ_* だけであり、トランジットライトカーブ解析から恒星の平均密度という物理量が分かることが示される。

4.2 透過光分光スペクトル

透過光分光はトランジット深さを波長方向に展開して測定する。波長依存性を含めたトランジット深さは

$$\delta(\lambda) = \left(\frac{R_p(\lambda)}{R_*} \right)^2 \quad (4.15)$$

のように表される。ここでは恒星半径の波長依存性は無視している。ここに $R_p(\lambda)$ は、光が透過しなくなる大気の高さに対応する半径となる。この高さは大気中の原子・分子吸収や散乱もしくは固体表面で決まる。図4.4に固体表面を持つ惑星に吸収のある分子を持つ大気がある場合の透過光分光の例を示す。分子吸収のない波長では、固体表面が R_p となる一方、吸収がある波長では大気中の光学的深さが1になる付近が R_p となるため、トランジット深さが少し深くなる。この違いを捉えることで大気中の分子を検出できる。この例では、分子吸収以外の波長の連続成分が固体表面で決まるとしたが、他の場合として、連続吸収やレイリー散乱、雲による散乱、強い吸収線のウイング、弱い

多数の吸収線の集合である pseudo continuum などもありうる。

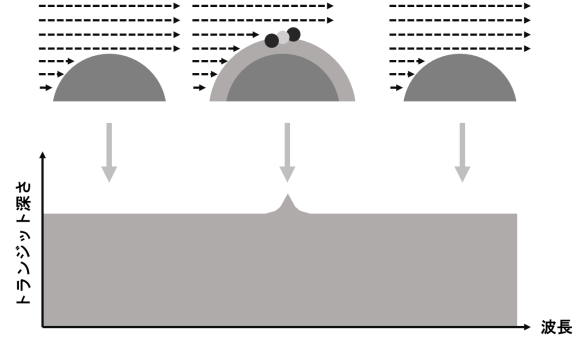


図4.4 透過光分光の仕組み。

4.3 惑星からの光

恒星光が惑星にやってくると一部は吸収され熱となり、一部は熱化されることなく宇宙に返される。前者は熱化されたエネルギーが輻射光として宇宙に返される。惑星が形成されて間もないと、惑星内部からのエネルギーもこれに加わる。後者は、光子をそのまま返す過程であり、一般的には反射・散乱と呼ばれる。散乱は大気中で確率的に光が進行方向を変更するが、反射は地表や雲などで離散的に進行方向を変更する。

系外惑星からの光を考える時、惑星が球体であることを適切に考慮しないとならない。輻射光は球体全体からの強度を考えねばならないし、さらに反射光は恒星と惑星と観測者の相対的な位置関係に左右される。

半径 R の惑星が受け取るエネルギーは、恒星フラックス密度 (stellar flux density) S を用いて

$$L_{\text{ab}} = (1 - A)\pi R^2 S \quad (4.16)$$

と書くことができる。 πR^2 は惑星の断面積である (図4.5参照)。 A は惑星全体の反射率 (アルベドという) で $(1 - A)$ の項は反射して宇宙に返された分を除いている。恒星光を黒体輻射と考えると、恒星フラックス密度は

$$S = \frac{L_*}{4\pi a^2} = \frac{4\pi R_*^2 \sigma T_*^4}{4\pi a^2} [\text{W/m}^2] \quad (4.17)$$

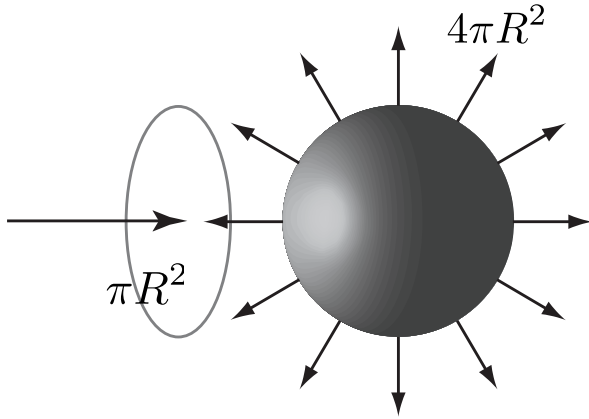


図4.5 惑星へ入射してくる恒星光と惑星表面から出ていく輻射光。

である。 a は公転半径、 R_* は恒星半径、 T_* は恒星温度、 σ はシュテファンボルツマン定数である。地球における $S_\odot = 1370 \text{ W/m}^2$ は太陽定数とよばれる。一方、惑星からの放射を温度 T_{eq} の黒体輻射とすると放射エネルギーは

$$L_{\text{em}} = \beta(4\pi R^2 \sigma T_{\text{eq}}^4) + L_{\text{int}} \quad (4.18)$$

とかける。ここで L_{int} は惑星自身のエネルギー放射である。若い惑星ではこの項が卓越しているため自己放射惑星と呼ばれる。

β は熱分配の度合いを表す係数で、瞬時に全球に一樣に分配する極限 $\beta = 1$ と、恒星光のあたっている半球だけに分配される場合 $\beta = 0.5$ の間の値をとる。潮汐力により公転と自転が一致（潮汐ロック）しているような恒星に近い惑星では、地球に対する月のように、惑星の同じ面が常に恒星方向を向くために導入される補正である。

通常の mature な惑星の場合、自己放射が無視できるため $L_{\text{int}} = 0$ とおける。この場合、恒星からの吸収エネルギーが輻射エネルギーと釣り合うはずである。

$$L_{\text{em}} = L_{\text{ab}} \quad (4.19)$$

とした時（放射平衡）の T_{eq} を放射平衡温度（単に平衡温度ともいう）と呼ぶ。この時、式 (4.16) から式

(4.19) までを用いると

$$\frac{T_{\text{eq}}^4 a^2}{T_*^4 R_*^2} = \frac{1-A}{4\beta} \quad (4.20)$$

となる。つまり

$$T_{\text{eq}} = \left(\frac{1-A}{4\beta} \right)^{\frac{1}{4}} T_* \sqrt{\frac{R_*}{a}} \quad (4.21)$$

$$= 396 \text{ K} \left(\frac{1-A}{4\beta} \right)^{\frac{1}{4}} \left(\frac{T_*}{5800 \text{ K}} \right) \left(\frac{R_*}{R_\odot} \right)^{\frac{1}{2}} \left(\frac{a}{1 \text{ au}} \right)^{-\frac{1}{2}} \quad (4.22)$$

$$= 396 \text{ K} \left(\frac{1-A}{4\beta} \right)^{\frac{1}{4}} \left(\frac{L_*}{L_\odot} \right)^{\frac{1}{4}} \left(\frac{a}{1 \text{ au}} \right)^{-\frac{1}{2}} \quad (4.23)$$

となる。

ここで惑星表面が受け取る平均的な恒星フラックスを定義しておこう。式 (4.16) より

$$F_* \equiv \frac{L_{\text{ab}}}{4\pi R^2} = \frac{(1-A)S}{4} \quad (4.24)$$

$$= 240 \left(\frac{1-A}{0.7} \right) \left(\frac{S}{S_\odot} \right) \text{ W/m}^2 \quad (4.25)$$

のように S と結びつく。この式は、恒星の入射エネルギーから惑星の反射分を取り除き、図4.5に示すような断面円で入ってきた入射フラックスが自転により球面へ分配される効果を表す係数 $1/4$ をかけたものになるという描像を表している。惑星が受け取る平均フラックス密度と放射平衡温度には、

$$\sigma T_{\text{eq}}^4 = F_*/\beta \quad (4.26)$$

の関係にあることもわかる。

Radiance, Irradiance

惑星や恒星は遠方から見れば点であり、近づいてみれば通常、連続的に密度が変わる三次元構造を持っている。しかし表面上のある基準球面、たとえば地球なら惑星表面や大気上端の高度一定面、恒星なら光球面などを考えると、ある面からの放射や、ある面への放射を考える事で、計算が易しくなる。そこで、まず、ある有限面積の表面からの放射や、表面への放射を定量化するための量を定義しよう。微小面積 dA からある方向に $d\Omega$ のコーン内（図4.6左）を通過する単位時間・単位波長あたりのエネルギー dE を考える。n

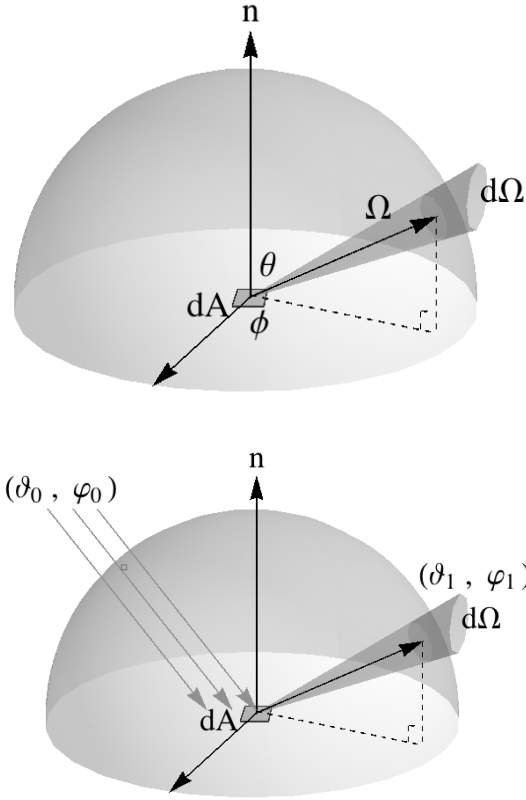


図4.6 (上) 素片 dA からの放射または、素片への放射。(下) (ϑ_0, φ_0) 方向からの平行光の入射と立体角 $d\Omega$ あたりの (ϑ_0, φ_0) 方向への反射。

は単位法線ベクトルである。 dE 自体は、見かけの投影面積 $\cos \theta dA$ に比例するので、

$$dE = L_{\uparrow}(\theta, \phi) \cos \theta dA d\Omega d\nu \quad (4.27)$$

のように比例定数を、radiance (放射輝度) L_{\uparrow} として定義すれば、投影の効果を除いて放射を定義できる。単位は例えば $[\text{W}/\text{m}^2/\text{sr}/\mu\text{m}]$ となる。 dA から \mathbf{n} 方向に流れるネットのエネルギーを irradiance (放射照度) E_{\uparrow} といい

$$E_{\uparrow} = \int_{\text{us}} dE = \int d\Omega L_{\uparrow}(\theta, \phi) \cos \theta \quad (4.28)$$

$$= \int_0^{2\pi} d\phi \int_0^{\pi/2} d\theta L_{\uparrow}(\theta, \phi) \cos \theta \sin \theta \quad (4.29)$$

となる。us は上半球面の意味である。

4.4 輻射スペクトル

半径 R_p で温度 T の球から発する黒体放射を距離 d で測定した時のフラックスを考える。温度 T を持った表面は黒体放射を行う。黒体放射の Radiance は

$$L_{\uparrow} d\lambda = B_{\lambda}(T) d\lambda = \frac{2hc^2}{\lambda^5} \frac{1}{\exp(hc/\lambda k_B T) - 1} d\lambda \quad (4.30)$$

である。素片 dA から $d\Omega$ の放射コーンを考え、距離 d にある面積 dA_{tel} の望遠鏡がコーンの先端と考える ($d\Omega = dA_{\text{tel}}/d^2$) と、 dA から dA_{tel} が受け取るエネルギー $\Delta E dA_{\text{tel}}$ は、

$$\Delta E dA_{\text{tel}} = L_{\uparrow} \cos \vartheta_1 d\Omega dA \quad (4.31)$$

$$= \frac{L_{\uparrow}}{d^2} \cos \vartheta_1 dA dA_{\text{tel}} \quad (4.32)$$

となる。すなわち、観測者から見た素片 dA による Irradiance もしくはフラックスは

$$\Delta E = \frac{L_{\uparrow}}{d^2} \cos \vartheta_1 dA \quad (4.33)$$

となる。

これを全球で積分して、

$$\begin{aligned} f_p &= \int_{\text{planet}} \Delta E = \int_{\text{planet}} dA \frac{B_{\nu}(T)}{d^2} \cos \vartheta_1 \quad (4.34) \\ &= R_p^2 B_{\nu}(T) \int_0^{2\pi} d\varphi_1 \int_0^{\pi/2} d\vartheta_1 \sin \vartheta_1 \cos \vartheta_1 \quad (4.35) \end{aligned}$$

$$= \pi B_{\nu}(T) \frac{R_p^2}{d^2} \quad (4.36)$$

となる。全エネルギーは波長方向の積分と距離 d の球殻面積をかけて

$$L = 4\pi d^2 \int_0^{\infty} d\nu \pi B_{\nu}(T) \frac{R_p^2}{d^2} = 4\pi R_p^2 \sigma T^4 \quad (4.37)$$

となる。

惑星からの輻射スペクトルは第ゼロ近似では、単一温度 T_p の黒体放射

$$f_p(\lambda) d\lambda = \pi B_{\lambda}(\lambda, T_p) \frac{R_p^2}{d^2} d\lambda \quad (4.38)$$

$$= \frac{2\pi hc^2}{\lambda^5} \frac{R_p^2}{d^2} \left[\exp\left(\frac{hc}{\lambda k_B T_p}\right) - 1 \right]^{-1} d\lambda, \quad (4.39)$$

で近似することができる場合がある。

例えば、ほぼ大気のない地球の放射スペクトルは、表面温度の $T = 200 - 300\text{K}$ の黒体放射でおおまかに近似できる。しかし、ガス惑星や褐色矮星に関しては、黒体放射の近似はあまりよくない。これは大気中の分子の強い吸収によるところが大きい。放射スペクトルは、大気中の光学的厚さが 1 付近になる深さの大気層の黒体放射に近い光によるが、これは大気中の分子により波長ごとに大きく変化する。この黒体放射からの大きな破れが、系外惑星大気の特徴である。

上向き射出フラックス

大気の放射平衡モデルでは、大気高さ方向の一次元モデルのフラックスを考えるが、境界条件に地表のフラックスを与えることがある。同様に黒体放射の素片から上向き射出されるフラックスを考えよう。素片周囲の上半面での角度積分を考えて

$$F_\nu(T) = \Delta E = \int_{\text{us}} d\Omega B_\nu(T) \cos \vartheta_1 = \pi B_\nu(T) \quad (4.40)$$

となる。全エネルギーでは同様に波長積分をして

$$F(T) = \int d\nu f_\nu(T) = \sigma T^4 \quad (4.41)$$

となる。

4.5 反射・散乱スペクトル

反射光は、反射面・観測者との相対位置により観測される強度がことなるので若干複雑である。しかし平均的には以下のように考えることができる。まず、式 (4.16) の時に、惑星が受け取らなかったエネルギーが反射となるので、反射のエネルギーは

$$L_p^{\text{ref}} = A\pi R_p^2 S = \frac{L_\star}{4\pi a^2} \pi R_p^2 A \quad (4.42)$$

である。距離 d で受け取る反射フラックスは平均的には

$$\langle f_p^{\text{ref}} \rangle = \frac{L_p^{\text{ref}}}{4\pi d^2} \quad (4.43)$$

であるから、恒星のフラックス

$$f_\star = \frac{L_\star}{4\pi d^2} \quad (4.44)$$

を用いて、

$$\langle f_p^{\text{ref}} \rangle = \frac{A}{4} \left(\frac{R_p}{a} \right)^2 f_\star \quad (4.45)$$

となる。

スペクトルの観点からは、波長の関数となっているのは A と f_\star の掛け算、つまり恒星スペクトルに反射スペクトルがかかったものであるという点である。

$$\langle f_p^{\text{ref}} \rangle(\lambda) d\lambda = \frac{1}{4} \left(\frac{R_p}{a} \right)^2 A(\lambda) f_\star(\lambda) d\lambda \quad (4.46)$$

である。

平均ではなく位相の関数としてフラックスを求めるのは少し複雑であるのでここでは割愛する。詳しくは拙著、系外惑星探査 5.1.1 を参照のこと。ここでは結果だけ紹介する。距離 d の位置から観測した惑星の反射光フラックス $f_p^{\text{ref}}(\lambda)$ は、主星-惑星間距離 a 、惑星アルベド $A(\lambda)$ 、惑星半径 R_p 、距離 d から観測した主星フラックス $f_\star(\lambda)$ 、また観測者から見た惑星位置による関数 $\phi(\beta)$ を用いて

$$f_p^{\text{ref}}(\lambda) = \frac{2\phi(\beta)}{3} A(\lambda) \left(\frac{R_p}{a} \right)^2 f_\star(\lambda), \quad (4.47)$$

$$\phi(\beta) \equiv [\sin \beta + (\pi - \beta) \cos \beta] / \pi, \quad (4.48)$$

と表される。ここで $\phi(\beta)$ は Lambert phase function とよばれ、位相角 (phase angle) $\beta = \angle(\text{主星}-\text{惑星}-\text{観測者})$ の関数である。ただしこの関係は、等方散乱を仮定していることに注意が必要である。海洋による鏡面反射 (ocean glint) などの非等方性の強い反射では必ずしも成り立たない。

恒星光が惑星にあたって反射・散乱する光は、現在の技術では、宇宙からの精密測光での位相カーブとしての検出などの限定された条件でしか達成されていないが、原理的には惑星表面の二次元分布情報を持つ豊かな情報源である。反射光で重要な指標は主星と惑星のフラックス比、主星惑星コントラスト

$$c_{\text{sp}}(\lambda) \equiv \frac{f_p(\lambda)}{f_\star(\lambda)} \quad (4.49)$$

である。主星惑星コントラストが緩いほうが直接撮像においても位相カーブにおいても検出が容易となる。

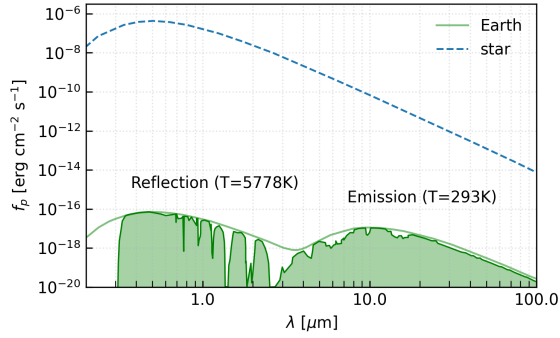


図4.7 地球のスペクトルとコントラスト

半月ならぬ半惑星 ($\beta = 90^\circ$) の場合、式 (4.47) を用いて、主星惑星コントラストを見積もることができる。地球の場合

$$c_{\text{sp}} \approx 10^{-10} \left(\frac{A}{0.3} \right) \left(\frac{R_p}{R_\oplus} \right)^2 \left(\frac{a}{1 \text{ au}} \right)^{-2} \quad (4.50)$$

程度に、ホットジュピターの場合

$$c_{\text{sp}} \approx 10^{-6} \left(\frac{A}{0.1} \right) \left(\frac{R_p}{R_J} \right)^2 \left(\frac{a}{0.05 \text{ au}} \right)^{-2} \quad (4.51)$$

程度になることがわかる。

4.6 様々な系外惑星スペクトル

図4.7は地球のスペクトルである。地球の場合、大気が薄いので輻射光はほぼ放射平衡温度に近い黒体となっている。また反射スペクトルは恒星スペクトルに大気吸収がかかったものとなっている。

図4.8は褐色矮星のスペクトル。黒体から大きく外れている。

図4.9は JWST NIRSPEC によるホットサターン WASP-39b のスペクトルである。分子吸収による feature が見て取れる。

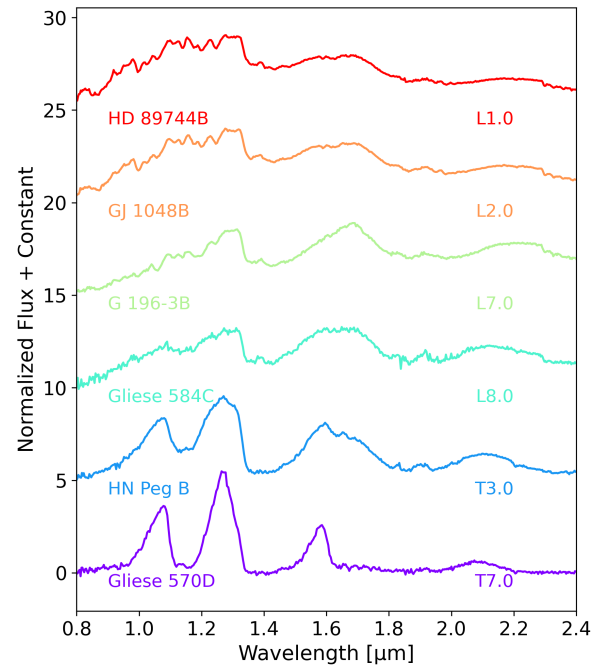


図4.8 褐色矮星のスペクトル (SpeX Prism ライブラリー) Inspired by[10]

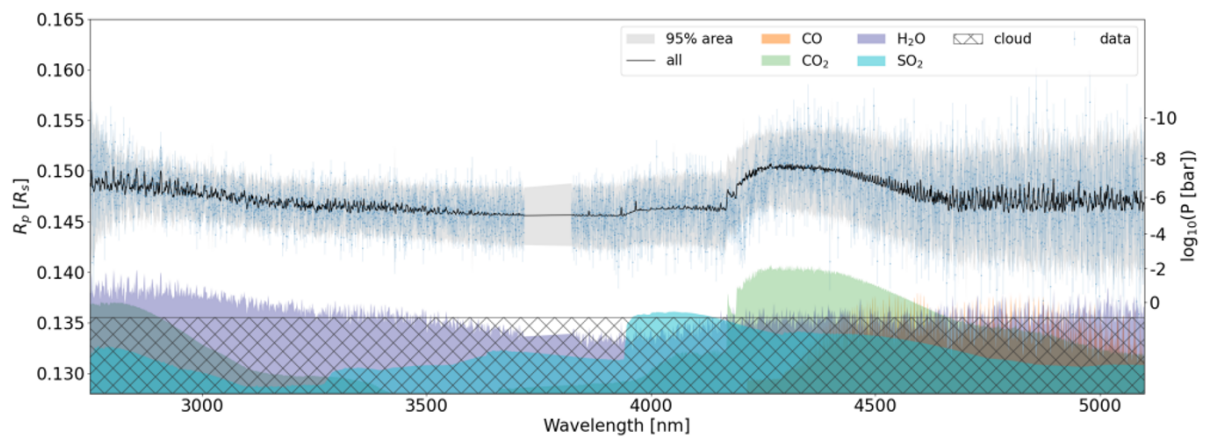


図4.9 JWST 透過光スペクトル WASP-39 b [5]

第 5 章

系外惑星大気

5.1 熱化学平衡における分子混合率

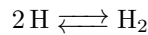
系外惑星大気中の分子の存在量は、高温高压の場合は熱化学平衡に近づくと考えられる。鉛直輸送や光化学反応、もしくは低温系大気の場合であっても、熱化学平衡の時にどうなるかは一つの基準として知っておくべきである。そこで、ここでは熱化学平衡に至っている場合の大気中の分子中の混合率がどのように決定されるか考える。

元素の保存

ガス中には H_2 や H 、 H_2O をいった分子種や原子種、イオン、自由電子が適当な割合で混合している。これらをまとめて種 (species) と呼ぶ。これら大気種中に含まれる原子の総数を考えたい。ガス中に含まれる種の中の原子 (例えば H , O , C) ではなく、ガス中のすべての分子や原子に含まれる原子を元素 (elements) と呼び区別する。本文書では元素をサンセリフ体を用いて、 H , O , C のようにあらわす。

二成分系の場合

まず最も単純な以下の水素原子と水素分子のガス 2 成分系



の場合、熱化学平衡でどのような組成比になるかを考えてみよう^{*1}。この場合、種は H と H_2 の二種類

^{*1} 実際は少量の触媒 M を介して $2H + M \rightleftharpoons H_2 + M$ としたほうが現実的だが、ここでは数学的側面を重視し M を無視する。

($N_s = 2$) であり、元素は H の一種類 ($N_e = 1$) である。

いま元素量として b_H (例えば 1 mol) の水素を考える。すなわち元素保存則は

$$n_H + 2n_{H_2} = b_H \quad (5.1)$$

となる。圧力 P と温度 T を与えたときに、熱化学平衡条件を満たす場合に、この元素 H がどのように H と H_2 に分配されるかを与えるのが、ギブス自由エネルギーの最小化である。すなわち熱化学平衡においては種の存在量は以下の最小化で与えられる

$$(n_H^*, n_{H_2}^*) = \text{minimize}_{(n_H, n_{H_2})} G(T, P, n_H, n_{H_2}) \quad (5.2)$$

$$\text{subject to } n_H + 2n_{H_2} = b_H \quad (5.3)$$

$$G(T, P, n_H, n_{H_2}) = n_H \mu_H + n_{H_2} \mu_{H_2} \quad (5.4)$$

$$n_H \geq 0, n_{H_2} \geq 0 \quad (5.5)$$

ここに μ_H 、 μ_{H_2} は H 、 H_2 の化学ポテンシャルであり、標準状態の化学ポテンシャルを用いると

$$\mu_H = \mu_H^o(T) + RT \log \frac{P_H}{P_{\text{ref}}} \quad (5.6)$$

$$= \mu_H^o(T) + RT \log \frac{n_H P}{(n_H + n_{H_2}) P_{\text{ref}}} \quad (5.7)$$

$$\mu_{H_2} = \mu_{H_2}^o(T) + RT \log \frac{P_{H_2}}{P_{\text{ref}}} \quad (5.8)$$

$$= \mu_{H_2}^o(T) + RT \log \frac{n_{H_2} P}{(n_H + n_{H_2}) P_{\text{ref}}} \quad (5.9)$$

で与えられる。 $\partial G / \partial n_H = \mu_H$ 、 $\partial G / \partial n_{H_2} = \mu_{H_2}$ が満たされることを確認せよ。

ここでは後の一般化を考え、上の等式条件付き最適化をラグランジュ未定乗数法で解く。ラグランジュ未

定乗数法では、自由パラメタを $\mathbf{x} = (n_{\text{H}}, n_{\text{H}_2}, \lambda)^\top$ の3パラメタとして

$$\mathbf{x}_* = \text{minimize}_{\mathbf{x}} \mathcal{L}(T, P, \mathbf{x}) \quad (5.10)$$

$$\mathcal{L}(T, P, \mathbf{x}) \equiv G(T, P, \mathbf{x}) + \lambda(n_{\text{H}} + 2n_{\text{H}_2} - b_{\text{H}}) \quad (5.11)$$

$$n_{\text{H}} \geq 0, n_{\text{H}_2} \geq 0 \quad (5.12)$$

を解けばよい。最後の非負条件は後でチェックする。 $\mathcal{L}(T, P, \mathbf{x})$ を \mathbf{x} で偏微分したものが $\mathbf{0}$ であるとして、 $\mathbf{x} = \mathbf{x}_*$ を求める。すなわち

$$\frac{\partial \mathcal{L}(T, P, \mathbf{x})}{\partial \mathbf{x}} = \begin{pmatrix} \mu_{\text{H}} + \lambda \\ \mu_{\text{H}_2} + 2\lambda \\ n_{\text{H}} + 2n_{\text{H}_2} - b_{\text{H}} \end{pmatrix} \quad (5.13)$$

$$= \begin{pmatrix} \mu_{\text{H}}^{\circ}(T) + RT \log \frac{n_{\text{H}} P}{(n_{\text{H}} + n_{\text{H}_2}) P_{\text{ref}}} + \lambda \\ \mu_{\text{H}_2}^{\circ}(T) + RT \log \frac{n_{\text{H}_2} P}{(n_{\text{H}} + n_{\text{H}_2}) P_{\text{ref}}} + 2\lambda \\ n_{\text{H}} + 2n_{\text{H}_2} - b_{\text{H}} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (5.14)$$

である。第1, 2成分から λ について消去すると

$$\log \left(\frac{n_{\text{H}}^2}{n_{\text{H}_2}(n_{\text{H}} + n_{\text{H}_2})} \frac{P}{P_{\text{ref}}} \right) = -\frac{2\mu_{\text{H}}^{\circ} - \mu_{\text{H}_2}^{\circ}}{RT} \quad (5.15)$$

がえられる。第3成分(保存則) $n_{\text{H}} + 2n_{\text{H}_2} - b_{\text{H}} = 0$ をもちいて、 n_{H_2} を消去すると、

$$\frac{b_{\text{H}}^2 - n_{\text{H}}^2}{4n_{\text{H}}^2} = \frac{P}{P_{\text{ref}}} \exp \left(-\frac{\mu_{\text{H}_2}^{\circ} - 2\mu_{\text{H}}^{\circ}}{RT} \right) \equiv k \quad (5.16)$$

であり、 $n_{\text{H}} \geq 0$ より

$$\frac{n_{\text{H}}}{b_{\text{H}}} = \frac{1}{\sqrt{4k+1}} \quad (5.17)$$

また保存則より

$$\frac{n_{\text{H}_2}}{b_{\text{H}}} = \frac{1}{2} \left(1 - \frac{1}{\sqrt{4k+1}} \right) \quad (5.18)$$

となる。これらは水素元素 b_{H} あたりの分子量となっていることに注意する。つまり実務上は $b_{\text{H}} = 1$ (mol) ととって計算を進めてもよい。

また体積分率 (Volume Mixing Ratio) は

$$\text{VMR}(\text{H}) = \frac{n_{\text{H}}}{n_{\text{tot}}} = \frac{1}{2} \left(\sqrt{k^2 + 4k} - k \right) \quad (5.19)$$

$$\text{VMR}(\text{H}_2) = \frac{n_{\text{H}_2}}{n_{\text{tot}}} = \frac{1}{2} \left(2 + k - \sqrt{k^2 + 4k} \right) \quad (5.20)$$

となる。図5.1に混合率の温度依存性を図示した。

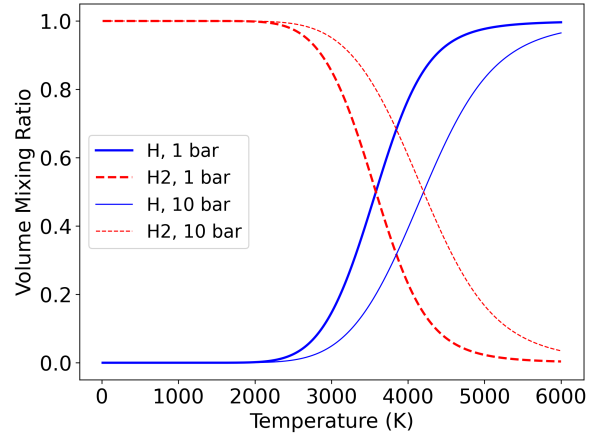
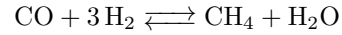


図5.1 $2\text{H} \rightleftharpoons \text{H}_2$ の熱化学平衡の体積混合率

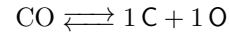
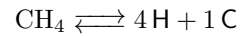
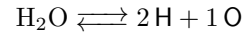
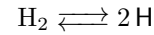
多成分系の場合[‡]

次に多元素系を考える。今、例として元素として $\text{H}, \text{C}, \text{O}$ 、また種としては $\text{CO}, \text{H}_2, \text{CH}_4, \text{H}_2\text{O}$ からなる系を考えよう。これらは水素大気のみ成分としては主要なものである。化学反応としては以下の一つの反応式となる。

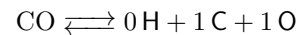
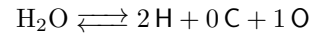
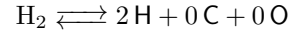


となる。

要素との化学反応式を考えると、



のようになる。ここで、0成分も含んで表現すると



のようになる。右辺について formula matrix を以下のように定義する。

$$A \equiv \begin{pmatrix} 2 & 2 & 4 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

種の存在量ベクトルを $\mathbf{n} = (n_{\text{H}_2}, n_{\text{H}_2\text{O}}, n_{\text{CH}_4}, n_{\text{CO}})^\top$, 元素の存在量ベクトルを $\mathbf{b} = (b_{\text{H}}, b_{\text{C}}, b_{\text{O}})^\top$ と置くと (元素の存在量を種とは区別している点に注意), 元素保存則は

$$A\mathbf{n} = \mathbf{b} \quad (5.21)$$

と表されることが分かる.

多元素系でも同様に, 以下の Gibbs エネルギーを最小化すればよい

$$\begin{aligned} \mathbf{n}^* &= \text{minimize}_{\mathbf{n}} G(T, p, \mathbf{n}) \\ \text{subject to } A\mathbf{n} &= \mathbf{b}, n_i \geq 0 \end{aligned} \quad (5.22)$$

$$G(T, p, \mathbf{n}) \equiv \boldsymbol{\mu}(T, P, \mathbf{n})^\top \mathbf{n} \quad (5.23)$$

$$\boldsymbol{\mu}(T, p, \mathbf{n}) = \boldsymbol{\mu}^\circ(T) + RT \log\left(\frac{p}{n_{\text{tot}}} \mathbf{n}\right) \quad (5.24)$$

ただし総数密度と規格化圧力を

$$n_{\text{tot}} = \sum_i n_i \quad (5.25)$$

$$p \equiv P/P_{\text{ref}} \quad (5.26)$$

と定義した.

$\text{CO}, \text{H}_2, \text{CH}_4, \text{H}_2\text{O}$ 系において, 式 (5.22) は,

$$\begin{aligned} \mathcal{L}(T, p, \mathbf{n}, \boldsymbol{\lambda}) &= \sum_{i=\text{CO}, \text{H}_2, \text{CH}_4, \text{H}_2\text{O}} \mu_i(T) n_i \\ &+ \lambda_{\text{H}}(2n_{\text{H}_2} + 4n_{\text{CH}_4} + 2n_{\text{H}_2\text{O}} - b_{\text{H}}) \\ &+ \lambda_{\text{C}}(n_{\text{CO}} + n_{\text{CH}_4} - b_{\text{C}}) \\ &+ \lambda_{\text{O}}(n_{\text{CO}} + n_{\text{H}_2\text{O}} - b_{\text{O}}) \end{aligned} \quad (5.27)$$

ここでは Gordon and McBride による NASA/CEA (Chemical Equilibrium with Applications) の実装法を元に解説する [3, 2]. CEA では Lagrange multiplier を用いて Gibbs エネルギー最小化を行う二次の手法の一つである. すなわち

$$(\mathbf{n}^*, \boldsymbol{\lambda}) = \text{minimize}_{(\mathbf{n}, \boldsymbol{\lambda})} \mathcal{L}(T, p, \mathbf{n}, \boldsymbol{\lambda}) \quad (5.28)$$

$$\mathcal{L}(T, p, \mathbf{n}, \boldsymbol{\lambda}) = G(T, p, \mathbf{n}) + \boldsymbol{\lambda}^\top (A\mathbf{n} - \mathbf{b}) \quad (5.29)$$

を最小化する. ただしこの時点では非負条件 $n_i \geq 0$ は外れていない.

$\mathcal{L}(T, p, \mathbf{n}, \boldsymbol{\lambda})$ の停留点を求めるために, 最適化変数のセット $\mathbf{y} = (\mathbf{n}, \boldsymbol{\lambda})$ で一次偏微分した関数を

$$\mathbf{f}(\mathbf{y}) \equiv \frac{\partial}{\partial \mathbf{y}} \mathcal{L}(T, p, \mathbf{y}) \quad (5.30)$$

を定義する. そして

$$\mathbf{f}(\mathbf{y}^*) = \mathbf{0} \quad (5.31)$$

の解を Newton 法でもとめることで, $\mathbf{y}^* = (\mathbf{n}^*, \boldsymbol{\lambda}^*)$ を求める. 次に式 (5.30) の \mathbf{n} 成分を計算する

$$\frac{\partial}{\partial \mathbf{n}} \mathcal{L}(T, p, \mathbf{n}, \boldsymbol{\lambda}) = \boldsymbol{\mu}(T, P, \mathbf{n}) + A^\top \boldsymbol{\lambda} \quad (5.32)$$

ここに熱力学より $\partial_{\mathbf{n}} G(T, p, \mathbf{n}) = \boldsymbol{\mu}(T, P, \mathbf{n})$ *2 および $\boldsymbol{\lambda}^\top (A\mathbf{n}) = (A^\top \boldsymbol{\lambda})^\top \mathbf{n}$ と $\partial_{\mathbf{x}}(S^\top \mathbf{x}) = S$ を利用した. 式 (5.32) は $\boldsymbol{\lambda}$ について一次である点に注意する. $\boldsymbol{\lambda}$ 成分は単に保存関係

$$\frac{\partial}{\partial \boldsymbol{\lambda}} \mathcal{L}(T, p, \mathbf{n}, \boldsymbol{\lambda}) = A\mathbf{n} - \mathbf{b} \quad (5.33)$$

である.

式 (5.24) を用いると Newton 法で解くべき方程式系 (5.31) は

$$\frac{\boldsymbol{\mu}^\circ(T)}{RT} + \log\left(\frac{p}{n_{\text{tot}}^*} \mathbf{n}^*\right) + \frac{A^\top \boldsymbol{\lambda}^*}{RT} = \mathbf{0} \quad (5.34)$$

$$A\mathbf{n}^* - \mathbf{b} = \mathbf{0} \quad (5.35)$$

となる. ただし n_{tot}^* は \mathbf{n}^* の関数, すなわち

$$n_{\text{tot}}^* = \sum_i n_i^* \quad (5.36)$$

となる.

ここで CEA のアルゴリズムではいくつかのトリックを用いる. まず, 式 (5.31) を求めるために, 独立変数に n_{tot} を追加する. もともと n_{tot} は式 (5.25) により \mathbf{n} に依存していたため, この条件, つまり式 (5.25) を, (前提条件ではなく) 制約条件として追加する. さらに \mathbf{n} , n_{tot} は対数を取ったものを独立変数に取り直す. つまり $\mathbf{q} = \ln \mathbf{n}$, $q_{\text{tot}} = \ln n_{\text{tot}}$ を独立変数とする. この操作により非負条件 $n_i \geq 0$ は自然に取り除かれる. また対数を取ることで広いダイナミックレンジで安定した数値計算となる. これにより新たな変数系として $\mathbf{z} = (\mathbf{q}, \boldsymbol{\lambda}, q_{\text{tot}})$ が導入され, 解くべき方

*2 式 (5.24) を用いて直接確かめることができる.

程式は

$$\mathbf{F}_n(\mathbf{z}) \equiv \frac{\mu^\circ(T)}{RT} + \mathbf{q} - q_{\text{tot}} \mathbf{u} + \log p \mathbf{u} + \frac{A^\top \boldsymbol{\lambda}}{RT} \quad (5.37)$$

$$\mathbf{F}_\lambda(\mathbf{z}) \equiv A e^{\mathbf{q}} - \mathbf{b} = \mathbf{0}_M \quad (5.38)$$

$$F_{\text{tot}}(\mathbf{z}) \equiv \sum_i e^{q_i} - e^{q_{\text{tot}}} = 0 \quad (5.39)$$

の解

$$\mathbf{F}_n(\mathbf{z}) = \mathbf{0}_M \quad (5.40)$$

$$\mathbf{F}_\lambda(\mathbf{z}) = \mathbf{0}_M \quad (5.41)$$

$$F_{\text{tot}}(\mathbf{z}) = 0 \quad (5.42)$$

である。 $\mathbf{0}_M$ は M 次元ゼロベクトルである。この解が $\mathbf{z}^* = (\mathbf{q}^*, \boldsymbol{\lambda}^*, q_{\text{tot}}^*)$ となる。これを簡潔に $\mathbf{F}(\mathbf{z}) = (\mathbf{F}_n(\mathbf{z}), \mathbf{F}_\lambda(\mathbf{z}), F_{\text{tot}}(\mathbf{z}))$ とまとめて扱うこともある。

Newton 法では $\mathbf{F}(\mathbf{z})$ を \mathbf{z} で微分したヤコビアンが必要となる。

$$\mathbf{J}(\mathbf{z}) = \begin{pmatrix} \frac{\partial \mathbf{F}_n}{\partial \mathbf{q}} & \frac{\partial \mathbf{F}_n}{\partial \boldsymbol{\lambda}} & \frac{\partial F_{\text{tot}}}{\partial q_{\text{tot}}} \\ \frac{\partial \mathbf{F}_\lambda}{\partial \mathbf{q}} & \frac{\partial \mathbf{F}_\lambda}{\partial \boldsymbol{\lambda}} & \frac{\partial F_{\text{tot}}}{\partial q_{\text{tot}}} \\ \frac{\partial F_{\text{tot}}}{\partial \mathbf{q}} & \frac{\partial F_{\text{tot}}}{\partial \boldsymbol{\lambda}} & \frac{\partial F_{\text{tot}}}{\partial q_{\text{tot}}} \end{pmatrix} = \begin{pmatrix} E_M & \frac{A^\top}{RT} & -\mathbf{u} \\ Y(\mathbf{q}) & \frac{Z_M}{RT} & \mathbf{0}_M \\ (e^{\mathbf{q}})^\top & \mathbf{0}_M^\top & -e^{q_{\text{tot}}} \end{pmatrix} \quad (5.43)$$

となる。ここに E_M は $M \times M$ 単位行列, Z_M は $M \times M$ ゼロ行列, M は元素ベクトル \mathbf{b} の要素数である。また行列 $Y(\mathbf{q})$ は $Y(\mathbf{q}) = A \text{diag}(e^{\mathbf{q}})$ のことであり, すなわち要素が

$$Y_{ij} = A_{ij} e^{q_j} \quad (5.44)$$

であるような行列である。Newton 法の update は式 (2.60) より, $\mathbf{J}(\mathbf{z}_k) \Delta \mathbf{z} = -\mathbf{F}(\mathbf{z}_k)$ をみたすので

$$\begin{pmatrix} E_M & \frac{A^\top}{RT} & -\mathbf{u} \\ Y(\mathbf{q}) & \frac{Z_M}{RT} & \mathbf{0}_M \\ (e^{\mathbf{q}_k})^\top & \mathbf{0}_M^\top & -e^{(q_{\text{tot}})_k} \end{pmatrix} \begin{pmatrix} \Delta \mathbf{q} \\ \Delta \boldsymbol{\lambda} \\ \Delta q_{\text{tot}} \end{pmatrix} = - \begin{pmatrix} \mathbf{F}_n(\mathbf{z}_k) \\ \mathbf{F}_\lambda(\mathbf{z}_k) \\ F_{\text{tot}}(\mathbf{z}_k) \end{pmatrix} \quad (5.45)$$

である。これより update $\Delta \mathbf{z} = (\Delta \mathbf{q}, \Delta \boldsymbol{\lambda}, \Delta q_{\text{tot}})$ は

$$\begin{aligned} \Delta \mathbf{q} + \frac{A^\top \Delta \boldsymbol{\lambda}}{RT} - \Delta q_{\text{tot}} \mathbf{u} \\ = -\frac{\mu^\circ(T)}{RT} - \mathbf{q}_k + (q_{\text{tot}})_k \mathbf{u} - \log p \mathbf{u} - \frac{A^\top \boldsymbol{\lambda}_k}{RT} \end{aligned} \quad (5.46)$$

$$Y(\mathbf{q}_k) \Delta \mathbf{q} = A(e^{\mathbf{q}_k} \odot \Delta \mathbf{q}) = \mathbf{b} - A e^{\mathbf{q}_k} \quad (5.47)$$

$$(e^{\mathbf{q}_k})^\top \Delta \mathbf{q} - e^{(q_{\text{tot}})_k} \Delta q_{\text{tot}} = e^{(q_{\text{tot}})_k} - \sum_i e^{(q_i)_k} \quad (5.48)$$

を満たす。

$\Delta \boldsymbol{\lambda}$ と $\boldsymbol{\lambda}_k$ が出てくるのは式 (5.46) のみである。我々は $\boldsymbol{\lambda}$ 自体には興味がないので, $\Delta \boldsymbol{\lambda}$ に代わる, あらたな update として

$$\boldsymbol{\pi} \equiv -\frac{\Delta \boldsymbol{\lambda} + \boldsymbol{\lambda}}{RT} \quad (5.49)$$

を定義し, $\mathbf{q}, \boldsymbol{\pi}, q_{\text{tot}}$ について update する。 $\boldsymbol{\pi}$ については使いきりであり保存しないことから Δ をつけない。式 (5.46) 式は

$$\Delta \mathbf{q} = A^\top \boldsymbol{\pi} + \Delta q_{\text{tot}} \mathbf{u} - \mathbf{g}_k(T) \quad (5.50)$$

$$\mathbf{g}_k(T) \equiv \frac{\mu^\circ(T)}{RT} + \mathbf{q}_k - (q_{\text{tot}})_k \mathbf{u} + \log p \mathbf{u} \quad (5.51)$$

と書けるため, $\boldsymbol{\pi}$ と Δq_{tot} が決まれば, 我々の求めたい $\Delta \mathbf{q}$ を算出することができる。つまり, 式 (5.50) を式 (5.47, 5.48) に代入した $M+1$ の連立方程式

$$\begin{aligned} A \text{diag}(e^{\mathbf{q}_k}) A^\top \boldsymbol{\pi} + \Delta q_{\text{tot}} A e^{\mathbf{q}_k} \\ = A(e^{\mathbf{q}_k} \odot \mathbf{g}_k(T)) + \mathbf{b} - A e^{\mathbf{q}_k} \end{aligned} \quad (5.52)$$

$$\begin{aligned} (A e^{\mathbf{q}_k})^\top \boldsymbol{\pi} + \Delta q_{\text{tot}} \left(\sum_i e^{(q_k)_i} - e^{(q_{\text{tot}})_k} \right) \\ = (e^{\mathbf{q}_k})^\top \mathbf{g}_k(T) + e^{(q_{\text{tot}})_k} - \sum_i e^{(q_i)_k} \end{aligned} \quad (5.53)$$

が解くべき方程式系となる。上の式では変数を $\mathbf{q}_k, (q_{\text{tot}})_k$ に保持したままの式にしたが, コーディング上は例えば $e^{\mathbf{q}}$ は \mathbf{n} に直しておいた方が見通しが良い。また,

$$\mathbf{B} \equiv A \text{diag}(\mathbf{n}_k) A^\top \quad (5.54)$$

$$\mathbf{b}_k \equiv A \mathbf{n}_k \quad (5.55)$$

$$\delta \mathbf{b}_k \equiv \mathbf{b} - \mathbf{b}_k \quad (5.56)$$

$$\delta n_{\text{tot},k} \equiv \sum_i n_{k,i} - (n_{\text{tot}})_k \quad (5.57)$$

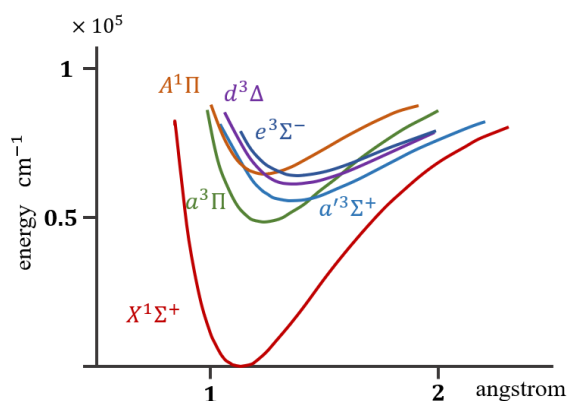


図5.2 一酸化炭素の電子状態。横軸は核子 (炭素-酸素) 間距離。

を定義する。ただし \mathbf{n}_k の i 成分を $n_{k,i}$ と表記した。これらを行うと

$$B\pi + \Delta q_{\text{tot}} \mathbf{b}_k = A(\mathbf{n}_k \odot \mathbf{g}_k(T)) + \delta \mathbf{b}_k \quad (5.58)$$

$$\mathbf{b}_k \cdot \pi + \delta n_{\text{tot},k} \Delta q_{\text{tot}} = \mathbf{n}_k \cdot \mathbf{g}_k(T) - \delta n_{\text{tot},k} \quad (5.59)$$

となる。この π と Δq_{tot} についての線形連立方程式を Reduced Gibbs Iteration Equations と呼ぶ。

5.2 二原子分子の振動・回転遷移

二原子分子では、電子準位間の遷移、核子の振動準位間の遷移、核子の回転準位間の遷移の組み合わせが量子力学的遷移となる。電子準位については、核子の運動が電子の運動に比べて極めて遅いため、核子間の距離 r を固定して電子の波動方程式を解き、そのエネルギー固有値を求めることができる。

核子間距離 r をゆっくり動かしたときに r に応じて基底状態の電子のエネルギー固有値が連続的に変化する。図5.2は一酸化炭素の場合の電子のエネルギー固有値（の一部）を核子間距離 r の関数として書いたものである。系外惑星大気中の分子吸収を考えるときは、通常、電子準位は基底状態（例えば CO なら $X^1\Sigma^+$ ）にあると考えてよい^{*3}。このため電子の基底状態のエネルギー固有値を r の関数として $V(r)$ と表

^{*3} 電子遷移を考える場合もある。

L	0	1	2	3	...
	S	P	D	F	...
縮重度	1	3	5	7	(2L+1)
Λ	0	1	2	3	...
	Σ	Π	Δ	Φ	...
縮重度	1	2	2	2	

表5.1 原子を考える場合の電子角運動量 L 、二原子分子を考える場合の分子軸 (z) 方向角運動量 L_z の絶対値 Λ と電子状態の記号の対応表。原子の場合、縮重度は $2L+1$ となるが、分子の場合、 $\Lambda=0$ を除き、 $L_z = -\Lambda, \Lambda$ の2種類である。

すことができ、核子 1,2 はポテンシャル $V(r)$ 中で運動しているとみなすことができる。

column – 電子状態のラベル[†]

二原子分子の場合、分子軸 (Z 軸) 周りの電子の合成軌道角運動量 L_z が保存することからこの絶対値 Λ を電子状態のラベルとする。 Λ の値に応じた分光学的記号は表5.1となる。原子の時と同様に分子でもスピン多重度が存在する。スピンを S ($S=0, 1/2, 1, 3/2, \dots$) とし多重度は $2S+1$ である。これを Λ の左上につけ

$$^{2S+1}\Lambda \quad (5.60)$$

の形で電子状態 n をラベルする。二原子分子が同じ原子二つからなっている場合 (等核二原子分子)、核子の位置の交換に対する操作に対し、波動関数の変化が $\Psi \rightarrow \pm\Psi$ の二種類の可能性がある。符号を変える場合 ungerade (奇数) の頭文字 u を、買えない場合 gerade (偶数) の頭文字 g を用いてこれを右下につける。さらに Σ 状態については分子軸を含む平面上の鏡像反転に関する波動関数の符号の変化の \pm を右上につけ、最終的に例えば

$$^2\Sigma_u^+ \quad (5.61)$$

のような記号となる。表5.2に、主な二原子分子の基底状態を示す。また、同じスピン多重度を持つ電子状態で、上のラベルの左側にエネルギーの低い順から X, B, C, D, \dots (基底状態を含む場合)、また基底状態と異なるスピン多重度に対しては、エネルギーの低い順に b, c, d, \dots を付ける表記もある。例えば、 $X^1\Sigma_g^+$ (基底状態)、 $B^1\Sigma_u^+$ 、 $C^1\Sigma_u^+$... $a^3\Sigma^-$...

大気中の一酸化炭素分子 (CO) のような、異なる質量 m_1, m_2 を持つ2つの核子と複数の電子から構成される自由に運動する二原子分子のエネルギー準位を考える。電子が常に基底状態にあるとすると、この系は、核子 1, 2 がポテンシャル $V(r)$ 中で運動して

分子	基底状態
H ₂	$1\Sigma_g^+$
N ₂	$1\Sigma_g^+$
O ₂	$3\Sigma_g^-$
CO	1Σ
CN	2Σ
NO	2Π

表5.2 代表的な二原子分子の電子基底状態

いるとみなすことができる．ここで r は核子間の距離である． $V(r)$ は図5.2の基底状態とすると、ある $r = r_e \sim 1.2$ で最小値をとり (V_e とする)、2つの核子は回転運動および $r = r_e$ を平衡距離とした振動運動を行う．

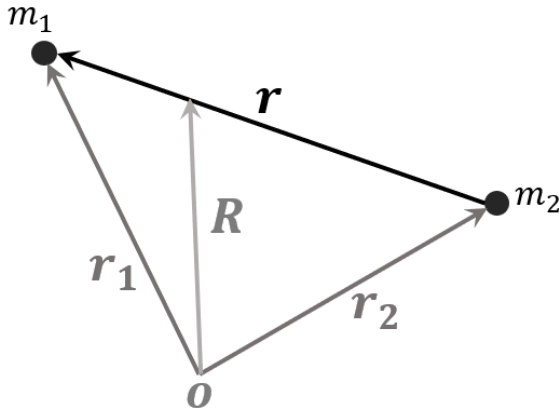


図5.3 二原子分子の核子の座標。

$V(r)$ が既知であるとして核子の波動方程式に着目する．図5.3のように核子の座標をとると、核子の波動方程式は、

$$\left(-\frac{\hbar^2}{2m_1} \nabla_1^2 - \frac{\hbar^2}{2m_2} \nabla_2^2 + V(r) \right) \psi(\mathbf{r}_1, \mathbf{r}_2) = E\psi(\mathbf{r}_1, \mathbf{r}_2) \quad (5.62)$$

と書ける． $\mathbf{r}_1, \mathbf{r}_2$ は核子 1, 2 の位置ベクトルであり、核子間の距離は $r = |\mathbf{r}_1 - \mathbf{r}_2|$ と表される．また ∇_1^2, ∇_2^2 は核子 1, 2 のラプラシアン、 \hbar は、プランク定数を h として $\hbar = \frac{h}{2\pi}$ 、 E は核子の波動方程式のエネルギー固有値、 $\psi(\mathbf{r}_1, \mathbf{r}_2)$ は核子の波動関数である．式 (5.62) を解くことで核子の振動・回転遷移を考える．

核子の重心座標 $\mathbf{R} = \frac{m_1 \mathbf{r}_1 + m_2 \mathbf{r}_2}{m_1 + m_2}$ と相対座標 $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ を定義し、 $\mathbf{r}_1, \mathbf{r}_2, \mathbf{R}$ 換算質量を $\mu = (m_1^{-1} + m_2^{-1})^{-1}$ 、総質量を $M = m_1 + m_2$ とすると重心座標・相対座標で書かれた波動方程式は

$$\left(-\frac{\hbar^2}{2\mu} \nabla_r^2 - \frac{\hbar^2}{2M} \nabla_R^2 + V(r) \right) \psi(\mathbf{r}, \mathbf{R}) = E\psi(\mathbf{r}, \mathbf{R}) \quad (5.63)$$

となる．ここに重心座標のラプラシアン $\nabla_R^2 = \frac{\partial^2}{\partial X^2} + \frac{\partial^2}{\partial Y^2} + \frac{\partial^2}{\partial Z^2}$ と相対座標のラプラシアン $\nabla_r^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ を定義した．

♣ 式 (5.63) †

\mathbf{R}, \mathbf{r} の三次元直交座標系の第一成分を X, x 、第二成分を Y, y 、第三成分を Z, z とする． x, X 座標の変換は

$$x = x_1 - x_2, X = \frac{m_1 x_1 + m_2 x_2}{M}$$

であるから

$$\frac{\partial}{\partial x_1} = \frac{\partial x}{\partial x_1} \frac{\partial}{\partial x} + \frac{\partial X}{\partial x_1} \frac{\partial}{\partial X} = \frac{\partial}{\partial x} + \frac{m_1}{M} \frac{\partial}{\partial X} \quad (5.64)$$

より

$$\frac{1}{m_1} \frac{\partial^2}{\partial x_1^2} = \frac{1}{m_1} \frac{\partial^2}{\partial x^2} + \frac{m_1}{M^2} \frac{\partial^2}{\partial X^2} + \frac{2}{M} \frac{\partial}{\partial x} \frac{\partial}{\partial X} \quad (5.65)$$

また

$$\frac{\partial}{\partial x_2} = \frac{\partial x}{\partial x_2} \frac{\partial}{\partial x} + \frac{\partial X}{\partial x_2} \frac{\partial}{\partial X} = -\frac{\partial}{\partial x} + \frac{m_2}{M} \frac{\partial}{\partial X} \quad (5.66)$$

より

$$\frac{1}{m_2} \frac{\partial^2}{\partial x_2^2} = \frac{1}{m_2} \frac{\partial^2}{\partial x^2} + \frac{m_2}{M^2} \frac{\partial^2}{\partial X^2} - \frac{2}{M} \frac{\partial}{\partial x} \frac{\partial}{\partial X} \quad (5.67)$$

(5.65, 5.67) より

$$\frac{1}{m_1} \frac{\partial^2}{\partial x_1^2} + \frac{1}{m_2} \frac{\partial^2}{\partial x_2^2} = \frac{1}{\mu} \frac{\partial^2}{\partial x^2} + \frac{1}{M} \frac{\partial^2}{\partial X^2} \quad (5.68)$$

同様に

$$\frac{1}{m_1} \frac{\partial^2}{\partial y_1^2} + \frac{1}{m_2} \frac{\partial^2}{\partial y_2^2} = \frac{1}{\mu} \frac{\partial^2}{\partial y^2} + \frac{1}{M} \frac{\partial^2}{\partial Y^2} \quad (5.69)$$

$$\frac{1}{m_1} \frac{\partial^2}{\partial z_1^2} + \frac{1}{m_2} \frac{\partial^2}{\partial z_2^2} = \frac{1}{\mu} \frac{\partial^2}{\partial z^2} + \frac{1}{M} \frac{\partial^2}{\partial Z^2} \quad (5.70)$$

を足して $-\hbar^2/2$ をかけると、

$$-\frac{\hbar^2}{2m_1} \nabla_1^2 - \frac{\hbar^2}{2m_2} \nabla_2^2 = -\frac{\hbar^2}{2\mu} \nabla_r^2 - \frac{\hbar^2}{2M} \nabla_R^2 \quad (5.71)$$

となる。

波動方程式 (5.63) に対し、波動関数 $\psi(\mathbf{R}, \mathbf{r})$ を重心

運動の波動関数 $\Phi(\mathbf{R})$ と相対運動の波動関数 $\phi(\mathbf{r})$ および重心運動のエネルギー固有値 E_R , 相対運動のエネルギー固有値 E_r を用いて, $\psi(\mathbf{R}, \mathbf{r}) = \Phi(\mathbf{R})\phi(\mathbf{r})$, $E = E_R + E_r$ と変数分離を行う. 重心運動の波動方程式と相対運動の波動方程式をそれぞれ求める。

$$\Phi(\mathbf{R}) \left(-\frac{\hbar^2}{2\mu} \nabla_{\mathbf{r}}^2 \phi(\mathbf{r}) + V(r)\phi(\mathbf{r}) - E_r \phi(\mathbf{r}) \right) + \phi(\mathbf{r}) \left(-\frac{\hbar^2}{2M} \nabla_{\mathbf{R}}^2 \Phi(\mathbf{R}) - E_R \Phi(\mathbf{R}) \right) = 0 \quad (5.72)$$

と変形できるので重心運動の波動方程式と相対運動の波動方程式としてそれぞれ

$$-\frac{\hbar^2}{2M} \nabla_{\mathbf{R}}^2 \Phi(\mathbf{R}) = E_R \Phi(\mathbf{R}) \quad (5.73)$$

$$-\frac{\hbar^2}{2\mu} \nabla_{\mathbf{r}}^2 \phi(\mathbf{r}) + V(r)\phi(\mathbf{r}) = E_r \phi(\mathbf{r}) \quad (5.74)$$

を得る。重心運動の波動方程式は分子自体の運動を表していて、例えば平面波解が対応する。

次に相対運動の波動方程式に対し、極座標 (r, φ, θ) を用いて

$$\phi(\mathbf{r}) = \frac{\phi_r(r)}{r} Y(\varphi, \theta) \quad (5.75)$$

のように変数分離を行う。 $\phi_r(r)$ に対する動径方向 r の波動方程式は、回転による遠心力項を $W(r)$ として、有効ポテンシャル $V_{\text{eff}}(r) = V(r) + W(r)$ 中を一次元運動する波動方程式に帰着する。 $W(r)$ を求めよう。

ラプラシアン の極座標表示は、 $\nabla_{r,\varphi,\theta}^2$ は、

$$\nabla_{r,\varphi,\theta}^2 = \frac{1}{r^2} \left(\frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \nabla_{\varphi,\theta}^2 \right) \quad (5.76)$$

$$\nabla_{\varphi,\theta}^2 = \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \quad (5.77)$$

で与えられ、球面調和関数 $Y_{lm}(\varphi, \theta)$ は

$$\nabla_{\varphi,\theta}^2 Y_{lm}(\varphi, \theta) + J(J+1) Y_{lm}(\varphi, \theta) = 0 \quad (5.78)$$

を満たす。ここで J は回転量子数 $J = 0, 1, 2, \dots$ であり、 m は $|m| \leq l$ を満たす整数である。

式 (5.75) を相対運動の波動方程式に代入し、ラプラシアンを極座標に変換すると

$$-\frac{\hbar^2}{2\mu} \left(\frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{1}{r^2} \nabla_{\varphi,\theta}^2 \right) \frac{\phi_r(r)}{r} Y(\varphi, \theta) + V(r) \frac{\phi_r(r)}{r} Y(\varphi, \theta) = E_r \frac{\phi_r(r)}{r} Y(\varphi, \theta) \quad (5.79)$$

となるが、これを変数分離すると

$$\begin{aligned} \frac{r^2}{\phi_r(r)} \frac{\partial^2}{\partial r^2} \phi_r(r) + \frac{2\mu r^2}{\hbar^2} (E_r - V(r)) \\ = -\frac{\nabla_{\varphi,\theta}^2 Y(\varphi, \theta)}{Y(\varphi, \theta)} \end{aligned} \quad (5.80)$$

となるから、これを $J(J+1)$ と置くことで $Y(\varphi, \theta) = Y_{lm}(\varphi, \theta)$ とおくことができる。

♣ 式 (5.80) †

$$\begin{aligned} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \frac{\phi_r}{r} \right) &= \frac{\partial}{\partial r} \left(r^2 \frac{r\phi_r' - \phi_r}{r^2} \right) \\ &= \frac{\partial}{\partial r} (r\phi_r' - \phi_r) = \phi_r' + r\phi_r'' - \phi_r' \\ &= r \frac{\partial^2}{\partial r^2} \phi_r(r) \end{aligned} \quad (5.81)$$

すると動径方向の波動方程式は

$$\begin{aligned} -\frac{\hbar^2}{2\mu} \frac{\partial^2}{\partial r^2} \phi_r(r) + \left(V(r) + \frac{\hbar^2 J(J+1)}{2\mu r^2} \right) \phi_r(r) \\ = E_r \phi_r(r) \end{aligned} \quad (5.82)$$

が得られ、これは

$$V_{\text{eff}}(r) = V(r) + \frac{\hbar^2 J(J+1)}{2\mu r^2} \quad (5.83)$$

とした有効ポテンシャル中の一次元の運動とみなすことができる。よって

$$W(r) = \frac{\hbar^2 J(J+1)}{2\mu r^2} \quad (5.84)$$

となる。

相対運動のエネルギー固有値 E_r について、振動量子数 ν と回転量子数 J の組を用いて、量子状態を (ν, J) で表記し、始状態を $(\nu, J) = (\nu_i, J_i)$, 終状態を $(\nu, J) = (\nu_f, J_f)$ とする。一酸化炭素のような二原子分子でみられる、 $\Delta\nu = \nu_f - \nu_i = 1$, $\Delta J = J_f - J_i = \pm 1$ の遷移を考える。始状態、終状態の相対運動のエネルギー固有値をそれぞれ $E_r = E_{r,i}$, $E_r = E_{r,f}$ と表記し、遷移エネルギーを $\Delta E = E_{r,f} - E_{r,i}$ とする。

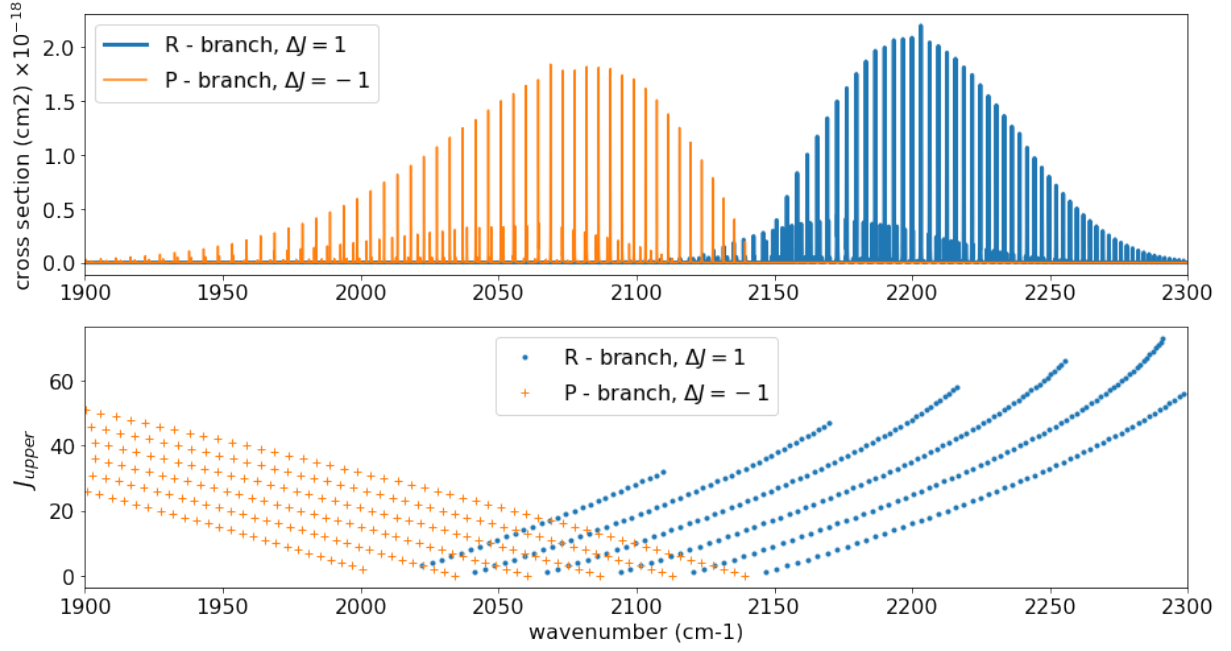


図5.4 一酸化炭素の $\Delta\nu = 1$ 遷移における R-branch, P-branch 断面積。

$q = r - r_e$ と座標変換する。

$$V_{\text{eff}}(q) \approx V_e + \frac{\mu}{2}\omega^2 q^2 + \frac{\hbar^2 J(J+1)}{2\mu(q+r_e)^2} \quad (5.85)$$

$$\approx V_e + \frac{\mu}{2}\omega^2 q^2 + \frac{\hbar^2 J(J+1)}{2\mu r_e^2} \quad (5.86)$$

となるから $\phi_r(x)$ に対する波動方程式は

$$-\frac{\hbar^2}{2\mu} \frac{\partial^2}{\partial q^2} \phi_r(q) + \frac{\mu}{2}\omega^2 q^2 \phi_r(q) = E'_r \phi_r(q) \quad (5.87)$$

$$E' = E_r - V_e - \frac{\hbar^2 J(J+1)}{2\mu r_e^2} \quad (5.88)$$

と調和振動子の波動方程式となる。角振動数 ω の調和振動子のエネルギー固有値は、 $\nu = 0, 1, 2, \dots$ を振動量子数として、

$$E' = \hbar\omega \left(\nu + \frac{1}{2} \right) \quad (5.89)$$

であることから、元の波動方程式のエネルギー固有値は

$$E_r = V_e + \hbar\omega \left(\nu + \frac{1}{2} \right) + \frac{\hbar^2 J(J+1)}{2\mu r_e^2} \quad (5.90)$$

となる。

$$E_r(\nu, J) = V_e + \hbar\omega \left(\nu + \frac{1}{2} \right) + \frac{\hbar^2 J(J+1)}{2\mu r_e^2} \quad (5.91)$$

とおく。 $E_{r,f} = E(\nu_f, J_f)$, $E_{r,i} = E(\nu_i, J_i)$ とおけるため

$$\Delta E = E_{r,f} - E_{r,i} = E(\nu_f, J_f) - E(\nu_i, J_i) \quad (5.92)$$

$$= \hbar(\nu_f - \nu_i) + \frac{\hbar^2}{2\mu r_e^2} [J_f(J_f + 1) - J_i(J_i + 1)] \quad (5.93)$$

$$= \hbar\Delta\nu + \frac{\hbar^2}{2\mu r_e^2} (J_f - J_i)(J_f + J_i + 1) \quad (5.94)$$

$$= \hbar\Delta\nu + \frac{\hbar^2}{2\mu r_e^2} \Delta J(2J_i + \Delta J + 1) \quad (5.95)$$

となる。

$\Delta\nu = 1$, $\Delta J = 1$ を式 (5.95) に代入すると

$$\Delta E = \hbar\omega + \frac{\hbar^2}{\mu r_e^2} (1 + J_i) \quad (5.96)$$

となる。 $\Delta\nu = 1$, $\Delta J = -1$ を式 (5.95) に代入すると

$$\Delta E = \hbar\omega - \frac{\hbar^2}{\mu r_e^2} J_i \quad (5.97)$$

となる。

これらの帰結は $\hbar\omega$ を中心に前後に間隔 $\frac{\hbar^2}{\mu r_e^2}$ で楕状に遷移が並ぶということである。図5.4は CO ($\Delta\nu = 1$) の断面積を示しており、楕状に並んでいる。波数で右側に伸びる式 (5.96) の系列を R-branch, 式 (5.97) を P-branch という。

5.3 ラインプロファイルと線強度

量子力学的効果により分子の遷移エネルギーは離散的であるが、吸収エネルギーは、その離散エネルギーの周りに様々な理由で広がる。この広がりを Broadening と呼ぶ。惑星大気では

- 熱運動や乱流によるドップラー広がり
- 不確定性原理による自然広がり
- 分子間のファンデルワールス力に由来する圧力依存の圧力広がり

の三種が重要な Broadening である。

まず、熱運動による broadening は、原子が熱運動により視線方向に速度 v_x を持つことでおこるドップラーシフト $\nu = \hat{\nu}(1 + v_x/c)$ に起因する。 v_x の分布関数は Maxwell 速度分布

$$P(v_x) = \sqrt{\frac{m}{2\pi k_B T}} e^{-\frac{mv_x^2}{2k_B T}} \quad (5.98)$$

$$= \sqrt{\frac{m}{2\pi k_B T}} e^{-\frac{mc^2(\nu - \hat{\nu})^2}{2k_B T \hat{\nu}^2}} \quad (5.99)$$

に比例してラインが広がること。HWHM (half width at half maximum),

$$\gamma_D = \hat{\nu} \sqrt{\frac{2(\log 2)k_B T}{mc^2}} \quad (5.100)$$

を用いて、ドップラー広がり line profile は

$$g_D(\nu; \hat{\nu}; \gamma_D) = \sqrt{\frac{\log 2}{\pi}} \frac{1}{\gamma_D} e^{-\log 2 \left(\frac{\nu - \hat{\nu}}{\gamma_D}\right)^2} \quad (5.101)$$

となる。すなわちドップラー広がりにはガウス分布に従う。乱流（ミクロ乱流）による速度分散も同様にガウ

ス分布で近似されるが、現在のところあまり考慮されていない。

一方、圧力広がりとは自然広がりと共に Lorentz profile,

$$g_L(\nu; \hat{\nu}; \gamma_L) = \frac{\gamma_L/\pi}{(\nu - \hat{\nu})^2 + \gamma_L^2} \quad (5.102)$$

で表される。ここに HWHM を γ_L としている。

惑星大気では特に圧力広がり (van der Waals Broadening) が重要である。地球大気の場合、 $p_0 = 1$ atmosphere, $T_0 = 296$ K での $\gamma_{L,W}$ を, air-broadening coefficient $\gamma_{L,W}^{\text{air}}$ といい、これを用いて、

$$\gamma_{L,W}(p, T) = \gamma_{L,W}^{\text{air}} \frac{p}{p_0} \left(\frac{T_0}{T}\right)^\alpha \quad (5.103)$$

のように圧力広がりを計算することが多い。ここに α は温度依存項のべき (temperature exponent) で代表的には 0.5 程度であるがさまざまである。圧力広がりには周囲の分子からのファンデルワールス力に起因するので、背景大気の種類によって異なる。特に系外惑星大気では水素分子・ヘリウム大気であることが多く、地球の大気下での値と異なることに注意が必要である。

自然広がりとは、不確定性原理からくるラインの広がりである。アインシュタイン A 係数を用いて

$$\gamma_{L,n} = \frac{A}{4\pi c} = \frac{0.222}{4\pi c} \left(\frac{\nu}{\text{cm}^{-1}}\right)^2 [\text{cm}^{-1}] \quad (5.104)$$

となる。これは電子が励起状態に滞在する確率が e^{-At} のようになることと不確定性関係に由来している。圧力広がりに対し、自然広がりには低圧下で寄与が大きくなる。

上記の Doppler profile と Lorentz profile の両方が効く場合、両者を convolution したものが line profile となる。これを Voigt profile と呼ぶ。

$$\begin{aligned} g_V(\nu; \hat{\nu}) &= (g_L * g_D)(\nu; \hat{\nu}) \\ &= \int_{-\infty}^{\infty} d\nu' g_L(\nu - \nu'; \hat{\nu}; \gamma_L) g_D(\nu' - \hat{\nu}; \hat{\nu}; \gamma_D) \end{aligned} \quad (5.105)$$

となる。圧力広がりとは自然広がり両方が寄与する場合は、 γ_L として

$$\gamma_L = \gamma_{L,W} + \gamma_{L,n} \quad (5.106)$$

を用いれば良い。これは Lorentian 同士の畳み込みが

$$\begin{aligned} g_L(\nu; \hat{\nu}, \gamma_{L,W}) * g_L(\nu; \hat{\nu}, \gamma_{L,n}) \\ = \int_{-\infty}^{\infty} d\nu' g_L(\nu - \nu'; \hat{\nu}; \gamma_{L,W}) g_L(\nu' - \hat{\nu}; \hat{\nu}; \gamma_{L,n}) \end{aligned} \quad (5.107)$$

$$\begin{aligned} &= \frac{(\gamma_{L,W} + \gamma_{L,n})/\pi}{(\nu - \hat{\nu})^2 + (\gamma_{L,W} + \gamma_{L,n})^2} \\ &= g_L(\nu; \hat{\nu}, \gamma_{L,W} + \gamma_{L,n}) \end{aligned} \quad (5.108)$$

となるからである。また Lorentian のテイルは無限に続くわけではなくある程度、線中心から離れると ($O(10^2)\text{cm}^{-1}$) 急激にカットオフが効くとされている。この Sub-Lorentian 効果についてはわかってないことが多いが、広い波数域で計算するには注意が必要である。

またラインプロファイルは

$$\int d\nu g(\nu) = 1 \quad (5.109)$$

に規格化されているため $g(\nu)$ の次元は cm である。

光子が分子によって吸収される過程を考える。分子分光では、エネルギーに対応する量を hc で割って、波数で表すことが一般的である。ここでもこの表記法に習おう。また、光の吸収を考えるので、終状態は始状態より高いエネルギーである。そこで終状態を up (upper state) は初期状態を low (lower state) のラベルに書き換える。例えば、 $E_{\text{low}} = E_{\nu_i, J_i}/hc$, $E_{\text{up}} = E_{\nu_f, J_f}/hc$ である。

光の吸収に関して各ラインの強度は、下の順位 E_{low} からの誘導吸収により生じるため、その状態の数に比例して強くなる。ただし、光が入射することで生じる誘導放射により実効的な吸収が減るため、この補正も必要である。

E_{low} の状態の数は、局所熱力学平衡にあると

$$\frac{n_{\text{low}}}{n} = \frac{g_{\text{low}}}{Q(T)} \exp\left(-\frac{hcE_{\text{low}}}{k_B T}\right) \quad (5.110)$$

となる。ここに g_{low} は縮退度、 n は全状態数、 $Q(T)$ は分配関数である。また誘導放射による補正は $(1 - e^{-hc\hat{\nu}/k_B T})$ である。ここに $\hat{\nu}$ は吸収波数、つまり upper/lower states の差の $\hat{\nu} = E_{\text{up}} - E_{\text{low}}$ のことで

ある*4。

係数も含めると線強度とはラインプロファイル $g(\nu)$ を用いて断面積が

$$\sigma^2(\nu) = S(T)g(\nu) \quad (5.111)$$

となる量である。つまり $S(T)$ の次元は cm である。

$$S(T) = \frac{g_{\text{up}}}{Q(T)} \frac{A}{8\pi c \hat{\nu}^2} e^{-c_2 E_{\text{low}}/T} \left(1 - e^{-c_2 \hat{\nu}/T}\right), \quad (5.112)$$

となる。ただし $c_2 \equiv hc/k_B = 1.4387773 \text{ cmK}$ である。 A はアインシュタイン A 係数である。

♣ 式 (5.112) †

実効的に起きる吸収は、誘導吸収 (B_{lu}) から誘導放射 (B_{ul}) を差し引いたものである。つまり

$$S(T) = \frac{hc\hat{\nu}}{c} \left(\frac{n_{\text{low}}}{n} B_{lu} - \frac{n_{\text{up}}}{n} B_{ul} \right) \quad (5.113)$$

が実効的に起こる吸収である。局所熱力学平衡の場合、

$$\frac{n_{\text{up}}}{n_{\text{low}}} = \frac{g_{\text{up}}}{g_{\text{low}}} \exp\left(-\frac{hc(E_{\text{up}} - E_{\text{low}})}{k_B T}\right) \quad (5.114)$$

$$= \frac{g_{\text{up}}}{g_{\text{low}}} e^{-hc\hat{\nu}/k_B T} \quad (5.115)$$

である。また詳細つり合いから

$$g_{\text{low}} B_{lu} = g_{\text{up}} B_{ul} \quad (5.116)$$

$$A = 2\pi hc \hat{\nu}^3 B_{ul} \quad (5.117)$$

となるが、これらを用いると

$$S(T) = \frac{h\hat{\nu}}{4\pi} \frac{n_{\text{low}}}{n} B_{lu} (1 - e^{-hc\hat{\nu}/k_B T}) \quad (5.118)$$

$$= \frac{g_{\text{up}}}{Q(T)} \frac{A}{8\pi c \hat{\nu}^2} e^{-hcE_{\text{low}}/k_B T} \left(1 - e^{-hc\hat{\nu}/k_B T}\right) \quad (5.119)$$

が得られる。途中で式 (5.110) を用いた。

5.4 様々な分子の振動回転遷移 †

メタン

図5.5に示すように、メタンの振動モードは4種類 $\nu_1, \nu_2, \nu_3, \nu_4$ があり、そのうち赤外活性つまり赤外域での光子吸収があるのは ν_3 と ν_4 である。これらは基本バンドとして、それぞれ $3.25\text{--}3.45 \mu\text{m}$, $7.5\text{--}8.0 \mu\text{m}$

*4 この量を E を使って表さないのは、分光的に問題となる実際に吸収中心である波数であるからであろう。

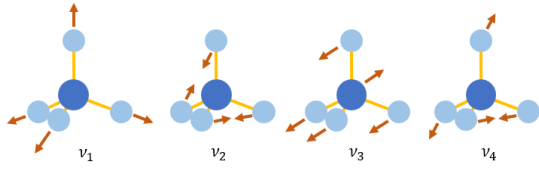


図5.5 メタンの4種振動モード

に強い吸収を持つ。またそれ以外の場所にも Polyad 構造による吸収（ホットバンド）が多数存在する。Polyad 構造は $\nu_1 \simeq \nu_3 \simeq 2\nu_2 \simeq 2\nu_4 \simeq 3000\text{cm}^{-1}$ となることに起因する構造であり、polyad 数

$$P = 2(\nu_1 + \nu_3) + \nu_2 + \nu_4 \quad (5.120)$$

で特徴づけられる。それぞれ Monad ($P=0$), Dyad ($P=1$), Pentad ($P=2$), Octad ($P=3$), Tetradecad ($P=4$), Icosad ($P=5$), Triacontad ($P=6$), Tetracontad ($P=7$), の名前がついていて、エネルギーは大まかに $1500P\text{cm}^{-1}$ 程度となる。例えば褐色矮星の H バンドに見られる $1.6\mu\text{m}$ 付近のメタン吸収は $P=4$ の Tetradecad に対応することが分かる [8]。

二原子分子のバンドヘッド

調和振動子近似ではポテンシャルが核子間距離の二次の関数であるとし、式 (5.85) の遠心力項が核子間の距離の関数であることを無視していた。実際、この近似は、図5.4のように、一酸化炭素の $\Delta\nu = 1$ 遷移をよく説明できる。しかし、振動遷移 $\Delta\nu$ が大きくなってきて核子間距離の、これらの近似が破れてきて、バンドヘッドという構造が表れる。ポテンシャルは

$$V_{\text{eff}}(q) \approx V_e + \frac{\mu}{2}\omega^2 q^2 + \frac{\hbar^2 J(J+1)}{2\mu r_e^2 [1 + (q/r_e)^2]} \quad (5.121)$$

$$\approx \frac{\hbar^2 J(J+1)}{2\mu r_e^2} \left[1 - 2\frac{q}{r_e} + 3\left(\frac{q}{r_e}\right)^2 \cdots \right] \quad (5.122)$$

と表されるが、この右辺第三項の q/r_e が大きくなると、遠心力が弱くなる効果が存在する。また、ポテンシャル自身も図5.2の基底状態で示すように平衡位置から離れると二次近似から逸脱してくる。これらの効

果を摂動論として計算すると、（省略するが）この計算は $X = (\nu + 1/2)$ と $Y = J(J+1)$ の二次式の形となる。

クロスターム (XY) 以外の二次の項は無視した形式を記すと

$$E_r(\nu, J) = V_e + \hbar\omega \left(\nu + \frac{1}{2} \right) + B_\nu J(J+1) \quad (5.123)$$

$$B_\nu = \frac{\hbar^2}{2\mu r_e^2} - \alpha_0 \left(\nu + \frac{1}{2} \right) \quad (5.124)$$

となる。これは振動が大きくなると、非調和ポテンシャルでは平均的な核間距離が延びることで、回転定数 B_ν が小さくなるという描像である。

式 (5.124) を用いて評価すると $B_{\nu+1} - B_\nu = -\alpha_0$ であるので、R, P-branch はそれぞれ

$$h\nu_{\text{line}}^R(J_u) = h\nu_\nu + 2BJ_u - \alpha_0 J_u^2 \quad (5.125)$$

$$h\nu_{\text{line}}^P(J_u) = h\nu_\nu - 2B(J_u + 1) - \alpha_0 (J_u + 1)^2 \quad (5.126)$$

$$B \equiv B_0 + \frac{1}{2}\alpha_0 \quad (5.127)$$

となり二次関数となる。

$\alpha_0 > 0$ とすると、R-branch の変曲点は $J_u = B/\alpha_0 > 0$ となるので*5、振動回転遷移の波数最大値（波長最小値）が存在し、このエネルギー付近に多数の遷移が集まることがわかる（図5.6の下パネル）。これを band head と呼ぶ。図5.6に 2.3 ミクロン付近の CO の band head の例を示した。今回横軸は波長であることに注意が必要である。band head 付近の準位が多い温度では、band head 付近の吸収が強くなる（図5.6の上パネル）。

また式 (5.125) と (5.126) を一つの式にまとめて、

$$h\nu_{\text{line}}(\mathcal{J}) = h\nu_\nu - 2B\mathcal{J} - \alpha_0 \mathcal{J}^2 \quad (5.128)$$

ただし、

$$\mathcal{J} = \begin{cases} J_u & (\mathcal{J} > 0) \\ -J_l & (\mathcal{J} < 0) \end{cases} \quad (5.129)$$

*5 一方、P-branch の変曲点は $-B/\alpha_0 < 0$ であるので、 $J_l = J_u + 1 > 0$ の領域に変曲点は存在しない。ただしこれは $\alpha_0 > 0$ の場合である。

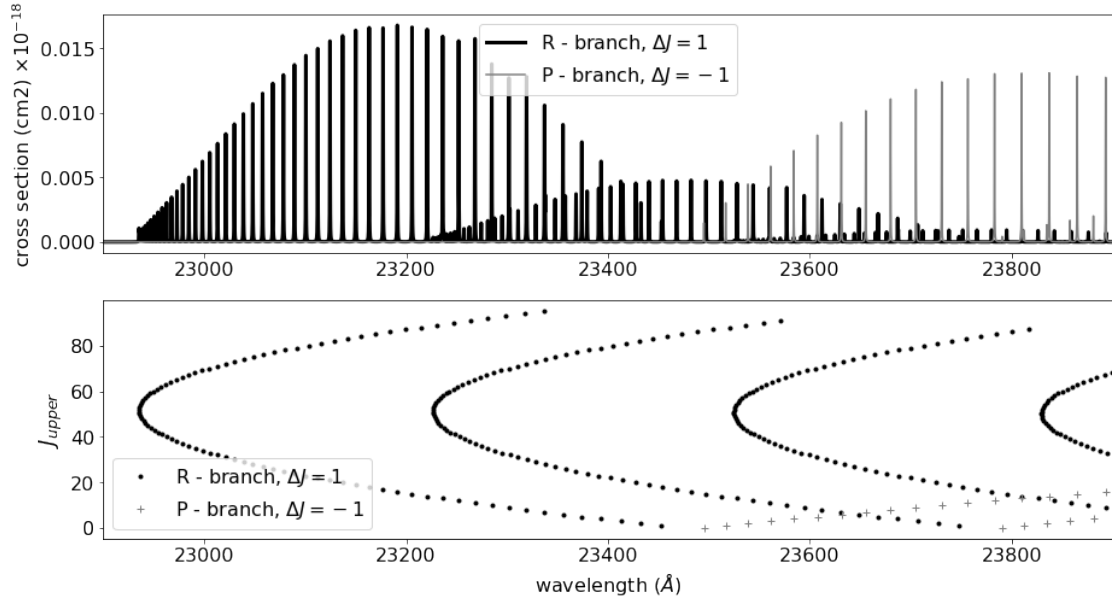


図5.6 CO 2.3 ミクロン帯 ($X^1\Sigma^+$, $\Delta\nu = 2$) の band head の例。 J_{upper} のエネルギー増加が減少に転じる点に多数の遷移が集まり band head が生じる。

としたものを、縦軸 J 、横軸ライン中心波数で書いたものを R. Fortrat にちなんでフォルトラ図という (図5.7)。振動エネルギーの差にあたる $J = 0$ の値 ν_v にはラインがない。これをバンド原点 (Band origin) と呼ぶ。

5.5 重力と惑星大気

惑星大気は、惑星の重力により惑星表面に束縛されている。すなわち重力が大気構造の大まかな構造をあたえる。大気構造に以下の仮定をおいて簡単化し、大気構造と重力の関係を見る。

- 大気層は薄く平衡平板で近似できる
- 大気は等温
- 大気は理想気体としてふるまう
- 大気の鉛直方向の運動はなく、重力と圧力が釣り合っている

ここから、大気重要な長さスケールであるスケールハイトが導入される。

理想気体の状態方程式

単一成分の理想気体の状態方程式は、圧力 P 、温度 T 、気体数密度 $n[\text{cm}^{-3}]$ を用いて

$$P = nk_B T \quad (5.130)$$

となる。アボガドロ数 $N_A = 6.0221367 \times 10^{23}$ を利用して

$$P = R'n'T \quad (5.131)$$

とも書ける。ここに $R' = N_A k_B = 8.3144598 \times 10^7 [\text{erg/K/mol}]$ は universal gas constant である。ここに n' はモル数密度 $[\text{mol cm}^{-3}]$ である。式 (5.130) を気体密度 $\rho = \mu m_H n [\text{g cm}^{-3}]$ (μ は分子量、 m_H は proton mass) であらわすと、

$$P = \frac{k_B}{\mu m_H} \rho T \quad (5.132)$$

$$(5.133)$$

となる。specific gas constant $R [\text{erg/g/K}]$ を用いてかくと

$$P = R\rho T \quad (5.134)$$

$$R \equiv \frac{k_B}{\mu m_H} \quad (5.135)$$

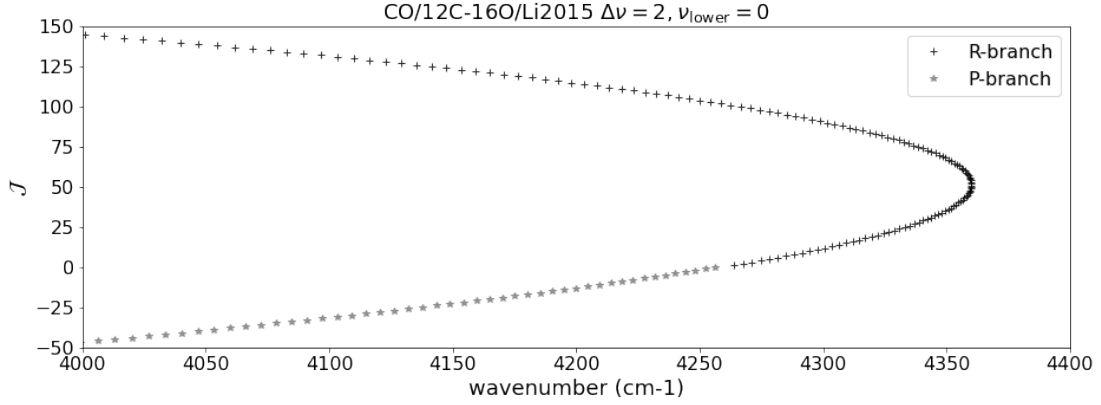


図5.7 CO のフォルトラ図。 $\Delta\nu = 2, \nu_{\text{lower}} = 0$ の場合を示している。

となる。以上のように R' を用いてモル数密度で考えているのか、 k_B もしくは R を用いて、通常の密度・数密度で考えているのか区別する必要がある。本稿では宇宙分野の表記との一貫性を保つため、原則は通常の密度・数密度で考える。しかし、気象学分野では数密度の mol 表記が一般的であり、比較の際や文献値を使用する際には表記の違いに注意が必要である。

等温・静水圧平衡

惑星の地表においた薄い大気層においては重力加速度

$$g = -\frac{d\phi}{dr} = \frac{GM_p}{r^2} \quad (5.136)$$

($\phi = GM_p/r$ は重力ポテンシャル) を r に寄らず一定と近似することができる。この条件下で静水圧平衡

$$\frac{dP(r)}{dr} = \rho \frac{d\phi}{dr} = -\rho g \quad (5.137)$$

に、状態方程式 (5.132) を用いると微分方程式

$$\frac{dP}{dr} = -\frac{P}{H} \quad (5.138)$$

の形となり解は

$$P(r) = P_0 \exp\left(-\frac{r-r_0}{H}\right) \equiv P_{\text{thin}}(r) \quad (5.139)$$

となる。 r_0 での圧力を P_0 として境界条件とした。ここに

$$H \equiv \frac{k_B T}{\mu m_H g} \quad (5.140)$$

$$\approx 8.4 \text{ km} \left(\frac{T}{300 \text{ K}}\right) \left(\frac{\mu}{30}\right)^{-1} \left(\frac{g}{980 \text{ cm/s}^2}\right)^{-1} \quad (5.141)$$

は (圧力) スケールハイトとよばれる。つまり、熱エネルギーと重力の比で大気の典型的な高さが決まる単純な描像が得られる。長さの次元を持つ大気の高さには熱エネルギー、すなわち温度情報が必要であることがわかる。

式 (5.140) から、例えば、温度の高い惑星のほうが大気の高さがあるため観測しやすいことなどがわかる。岩石惑星の場合、密度がほとんど半径によらないため、 H は半径に反比例する。すなわち半径が二倍地球半径のスーパー・アースは、半径は二倍になるが大気の厚さは半分になるため、透過分光による大気キャラクター化の難しさはあまり変わらない。また、式 (5.139) から、ある $r > r_0$ の r_0 からの高さを圧力から求めるには、

$$\Delta r = (r - r_0) = H \log\left(\frac{P_0}{P(r)}\right) \quad (5.142)$$

となる。

また大気中で等温とみなせる薄い一層を考え、そのレイヤーの高さ幅を $dz (> 0)$ 、圧力単位での幅を

$dP(>0)$ とすると、式 (5.140) から

$$\frac{dP}{P} = \frac{dz}{H} \quad (5.143)$$

と書ける。このようにスケールハイトで規格化された高さの変化が圧力の相対変化と一致する。改めて書くほどではないが、式 (5.137) から高さ座標と圧力座標の変換は

$$dz = \frac{dP}{\rho g} \quad (5.144)$$

である。

5.6 大気と分子存在量

前節では等温であるとしたが、一般には大気は等温ではない。大気の高さを長さの次元で考えるためには、上で見たように温度情報が必要であるが、これが等温でない場合、複雑になってしまう。そこで大気の高さを長さの次元ではなく、圧力で代用する方法がよく用いられる。この場合、温度の鉛直構造は縦軸圧力・横軸温度であらわされるが、圧力軸は鉛直方向を示すために逆転させることが多い。

また大気からの放射や透過・反射を考える場合、光学的厚さと圧力を変換する必要がある。式 (5.137) より、

$$dr = -\frac{dP}{\rho g} \quad (5.145)$$

であるから、断面積を σ とした光学的厚さの微分形は

$$d\tau = -n\sigma dr = \frac{\sigma}{\mu m_H g} dP \quad (5.146)$$

と表される。ただし符号は τ をどちらから測るかに依存する。今は、 $r \rightarrow \infty$ を $\tau \rightarrow 0$ と定義している。この表記でも陽に温度にはよらないことに注意。ただし、一般的に断面積は温度・圧力に依存する。またガス組成の鉛直分布依存に対応して平均分子量も厳密には圧力に依存する、すなわち

$$d\tau = \frac{\sigma(T, P)}{\mu(P) m_H g} dP \quad (5.147)$$

である。ただし、ここでは薄い大気を考えているので g は圧力によらないことに注意。

多成分系

大気が多成分系からなる場合はどうだろうか？ m_H をプロトン質量とすると、第 i 成分の分圧は

$$P_i = k_B n_i T = R \rho_i T \quad (5.148)$$

である。今、密度と数密度の関係は

$$\rho = \sum_{i=1}^N \rho_i = m_H \sum_{i=1}^N \mu_i n_i \quad (5.149)$$

$$= m_H \left(\sum_{i=1}^N \mu_i \frac{n_i}{n} \right) n = m_H \mu n \quad (5.150)$$

$$\mu \equiv \sum_{i=1}^N \xi_i \mu_i \quad (5.151)$$

と表せる。ここに総数密度 $n = \sum_{i=1}^N n_i$ を定義した。また μ は平均分子量であり、 $\xi_i = n_i/n$ は体積混合率 (Volume Mixing Ratio; VMR) と呼ばれる量である。体積混合比は分圧と

$$P_i = \xi_i P \quad (5.152)$$

の関係がある^{*6}。また式 (5.151) の両辺を μ で割ることにより

$$1 = \sum_{i=1}^N X_i \quad (5.153)$$

$$X_i = \frac{\mu_i}{\mu} \xi_i \quad (5.154)$$

$$= \frac{\rho_i}{\rho} \quad (5.155)$$

となる。 X_i は、全ガスに占める分子 i の質量の割合を示し、質量混合比 (Mass Mixing Ratio; MMR) と呼ばれる。

密度表記の状態方程式は平均分子量 μ を用いて

$$P = \sum_{i=1}^N P_i = k_B n T = R \rho T \quad (5.156)$$

$$R \equiv \frac{k_B}{\mu m_H} \quad (5.157)$$

と一成分系のように扱える。

^{*6} 体積混合比は、ガスの気体分子全体の数に対する分子 A の数の割合であり、なぜこの量を体積混合比と呼び、数混合比と叫ばないのかは不明である。

分子存在量推定と基本縮退

分子多成分系の場合、第 i 成分の opacity は、式 (5.152)、式 (5.154) より

$$d\tau_i = -n_i \sigma_i dr = \frac{\sigma_i}{\mu m_H g} dP_i \quad (5.158)$$

$$= \frac{\xi_i \sigma_i}{\mu m_H g} dP \quad (5.159)$$

$$= \frac{X_i \sigma_i}{\mu_i m_H g} dP \quad (5.160)$$

となる。

ここで式 (5.160) をよく見ると、 σ_i, μ_i, m_H は大気とは無関係に物理化学的に決まる量であるが、 X_i と g は系に固有の量であり、観測の観点からは推定量である。しかし $d\tau_i$ は X_i/g にのみ依存しているため、これを同時に決めることはできない。分子だけでオパシティが決まる場合の総光学的厚さは、線形和

$$d\tau = - \sum_{i=1}^N n_i \sigma_i dr = \sum_{i=1}^N d\tau_i \quad (5.161)$$

となっていることから、結局、スペクトルから推定できるのは、分子存在量ではなく、質量混合比を重力で割った X_i/g のみである。これを我々は基本縮退と呼んでいる。

トランジット系の場合、RV とトランジット半径から g が決まるため、基本縮退は解けるの問題はない。直接撮像惑星や褐色矮星の直接分光のような放射光のみの場合、photosphere が分子吸収しか見えない場合、基本縮退は解けない。しかし水素分子等による連続吸収が見える場合は基本縮退が解けるようになる [7]。

第 6 章

スペクトルからの大気推定

スペクトルから惑星大気の諸量を推定する手順は、大まかには以下のとおりである。

1. 大気理論モデルの設定：パラメータを持つ大気モデルを作成し、モデルからスペクトルを生成できるコードを作成する
2. 確率モデルの設定：スペクトルデータに付随する誤差モデルと大気モデルの事前分布を設定する
3. ベイズ推定の遂行：実スペクトルデータと生成したスペクトルから尤度を計算し、大気モデル中のパラメータの事後確率分布を求める

大気理論モデルは、最も物理化学過程を含まないフリーリトリvableでは、温度圧力分布自体を推定する。適当なパラメトライズモデルかガウス過程を用いたより自由なプロファイルを仮定することもできる。分子の存在量もパラメータとすることもできる。また高温系では熱化学平衡を仮定して、元素存在量を推定パラメータとすることも可能である。

大気理論モデルが設定されたら、次にスペクトルを生成する。以下ではこのスペクトル生成部分を詳述する。ベイズ推定の部分は第3章と基本的には同様である。

6.1 透過光スペクトルのモデリング

図6.1がトランジット時の概念図である。透過光で問題になるのは惑星半径付近の薄い大気層での透過率の波長依存性である。そこで、十分透過率が低い位置の基準半径 R_0 を定めてそこから上層までの大気層の円環成分 (図6.1C) の実効的な影面積 $A(\lambda)$ (以降、

円環影と呼ぶ) を考える。

例えばトランジット深さだけ考える場合、波長 λ で惑星半径と恒星半径を $R_p(\lambda), R_\star(\lambda)$ であるとして

$$\delta(\lambda) = \frac{R_p^2(\lambda)}{R_\star^2(\lambda)} = \frac{\pi R_0^2 + A(\lambda)}{\pi R_\star^2(\lambda)} \quad (6.1)$$

が観測量となり、円環影部分のみを考えればよい^{*1}。

さて一般に円環影部分 (図6.1C) の計算を考えよう。座標系を図6.1 B/D のように取る。円環影部分は

$$\begin{aligned} A &= \int_0^{2\pi} \int_{R_0}^{\infty} [1 - \mathcal{T}_\lambda(r_c, \phi)] r_c dr_c d\phi \\ &= \pi \left(2 \int_{R_0}^{\infty} [1 - \mathcal{T}_\lambda(r_c, \phi)] r_c dr_c \right) \end{aligned} \quad (6.2)$$

また、

$$R_p(\lambda) = \sqrt{R_0^2 + 2 \int_{R_0}^{\infty} [1 - \mathcal{T}_\lambda(r_c, \phi)] r_c dr_c} \quad (6.3)$$

とあらわされる。ここに $\mathcal{T}_\lambda(r_c, \phi)$ は、ある波長 λ における惑星ディスク半径 r_c での弦方向の透過率である。 $\mathcal{T}_\lambda(r)$ は弦方向の光学的厚さ t と

$$\mathcal{T}_\lambda(r_c) = e^{-t} \quad (6.4)$$

の関係にある。以降、表記の簡略化のため λ の添え字を省略する。

t は弦方向の座標を z と取り、

$$t(r_c) = \int_{-\infty}^{\infty} \kappa(r) \rho(r) dz = 2 \int_0^{\infty} \kappa(r) \rho(r) dz \quad (6.5)$$

^{*1} Ingress/Egress を考える場合は円環影の一部分しか透過光に寄与しないためより複雑となるがここでは考えない。

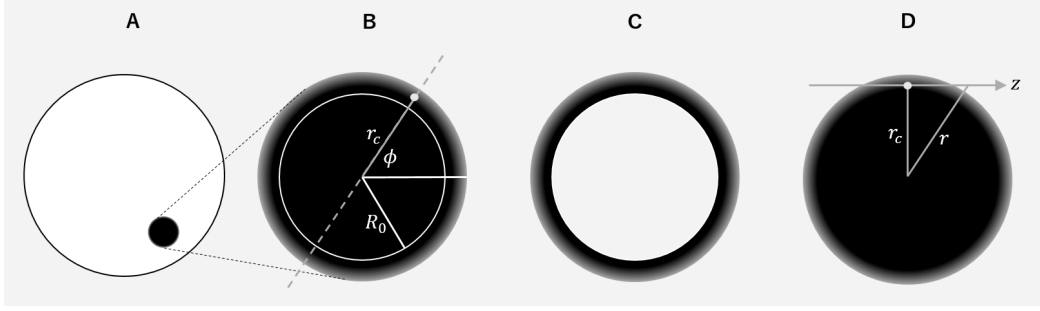


図6.1 透過分光を考えるための座標系。A：恒星の前面を通過する惑星。B：惑星部分の拡大図。C：円環影部分。D：中の破線方向に惑星を切断したときの図。恒星光の通過方向である z 方向を弦方向と呼んでいる。右図は「系外惑星探査」図 4.13 に対応している。

である。さてこの積分を大気レイヤーモデルで評価していこう。 t_n を n 番目のレイヤーの弦方向のオパシティとする。図6.2より

$$t_n = 2 \sum_k \kappa_k \rho_k \Delta z_k^{(n)} \quad (6.6)$$

この $\Delta t_k^{(n)}$ は n 番目のレイヤーにむかってはられた弦方向の $k - k - 1$ 間レイヤーのオパシティである。これを鉛直一次元方向のオパシティ $\Delta \tau_k = \kappa_k \rho_k \Delta r_k$ と関係づけると

$$t_n = \sum_k C_{nk} \Delta \tau_k \quad (6.7)$$

$$C_{nk} \equiv \frac{2 \Delta z_k^{(n)}}{\Delta r_k} = 2 \frac{\sqrt{r_{k-1}^2 - r_n^2} - \sqrt{r_k^2 - r_n^2}}{r_{k-1} - r_k} \quad (6.8)$$

となる。また $r_{N-1} = R_0$ である。ここに下三角行列である Chord Geometric matrix $C = \{C_{nk}\}$ を要素ごとに定義した。

N は総レイヤー数 ($n = 0, 1, \dots, N - 1$) とする。各レイヤー（下側境界までの）半径とレイヤーの高さには

$$r_{n-1} = r_n + \Delta h_n \quad (6.9)$$

の関係がある。境界条件は、大気レイヤーの最下層を $n = N - 1$ とし、 $R_0 = r_{N-1}$ を惑星中心から最下層の下側境界までの基準半径とする。 r_{N-1} より下は光が全く透過しないとする。すなわち

$$r_n = \begin{cases} \sum_{j=N-1}^{n+1} h_j + R_0 & \text{for } j < N - 1 \\ R_0 & \text{for } j = N - 1 \end{cases} \quad (6.10)$$

となる。また $n = 0$ のレイヤーの上側境界 (Top of Atmosphere; TOA) が必要である。これは別途 r_{top} と表記する。

大気レイヤーは (T_n, P_n) で表記されているため、長さスケールをスケールハイトを用いて変換しないといけない

$$\Delta h_n = H_n \frac{\Delta P_n}{P_n} \quad (6.11)$$

$$H_n = \frac{k_B T_n}{g_n \mu_n m_H} \quad (6.12)$$

また重力定数は

$$g_n = \frac{G(M_p + \Delta M_n)}{r_n^2} \quad (6.13)$$

と表される。 $\Delta M_n = \sum_{j=N-1}^n \Delta m_j$ であり m_j は j 番目のレイヤーの持つ質量である。 ΔM_n 自体は通常は M_p に比べて無視できるので考えなくても良いかもしれない。

数値的には最下層から順番に $n = N - 1, N - 2, \dots, 0$ について $\Delta h_n, r_n$ を解いていけばよい。境界条件は $r_N = R_0, g_N = GM_p/r_N^2$ である。また $h_0 = 0$ として最上層のオパシティは含めない。また同時に各層ごとに C_{nk} を $k = n$ から $k = N - 1$ まで計算をしておけば、Chord Geometric Matrix が計算できる*2。このように C は三角行列である。

*2 もちろん r_n を計算した後に式 (6.8) で一括で計算してもよい。

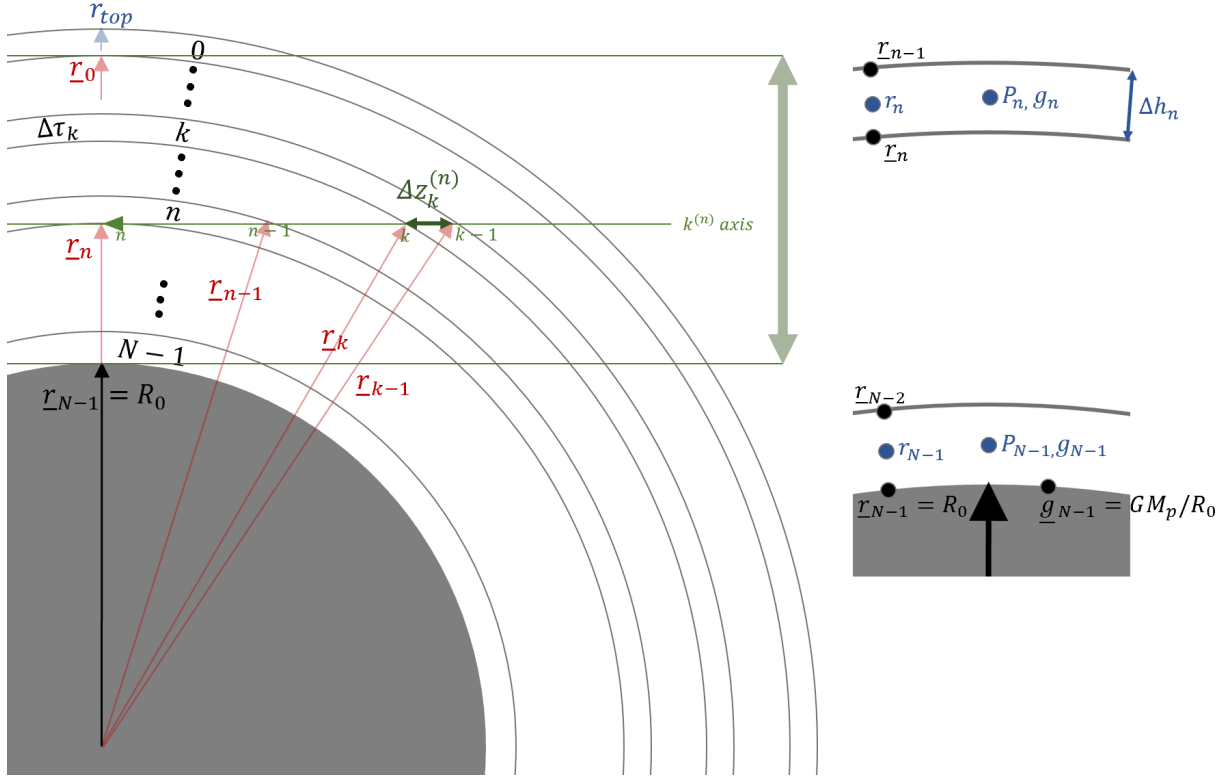


図6.2 レイヤーモデルにおける光学的厚さの計算方向。

6.2 Schwarzschild 方程式 *

ある角度 (\mathbf{n}) 方向に、ある波数幅 $d\nu$ の間の光が微小面積 (dS 、法線方向を \mathbf{k} とする)・微小時間、微小立体角 $d\Omega$ 内に伝達するエネルギーを $d\mathcal{E}_\nu$ とすると、specific intensity \mathcal{I}_ν は以下のように表現される

$$d\mathcal{E}_\nu = \mathcal{I}_\nu (\mathbf{n} \cdot \mathbf{k}) dS d\Omega d\nu dt. \quad (6.14)$$

微小円柱に入射した光は \mathcal{I}_ν に比例して散乱・吸収する。比例定数を extinction coefficient (減光の係数という意味) κ_ν という。ところで **opacity** という語は様々な定義で使われるが、この $[cm^2/g]$ の次元を持つ extinction coefficient を opacity と呼ぶことにする。また、今の意味では opacity は単一周波数 ν に対して定義されているので厳密には monochromatic opacity である。何らかの周波数平均をした代表的な extinction coefficient も opacity と呼ぶ。この opacity を用いると、微小円柱内からの射出がない場合、単位

時間内に吸収・散乱されるエネルギーは、微小距離 ds 、密度 ρ を用いて、

$$-\kappa_\nu \mathcal{I}_\nu \rho ds d\Omega d\nu dt = d\mathcal{I}_\nu d\Omega d\nu dt \quad (6.15)$$

となるので、

$$d\mathcal{I}_\nu = -\kappa_\nu \mathcal{I}_\nu \rho ds \quad (6.16)$$

となる。微小円柱内からの射出放射輝度は emission coefficient η_ν を用いて

$$d\mathcal{I}_\nu = \eta_\nu \rho ds \quad (6.17)$$

と定義するので、全体では

$$d\mathcal{I}_\nu = -\kappa_\nu \mathcal{I}_\nu \rho ds + \eta_\nu \rho ds \quad (6.18)$$

となる。ここで放射源関数 (source function)

$$\mathcal{J}_\nu \equiv \frac{\eta_\nu}{\kappa_\nu} \quad (6.19)$$

を定義すると、放射伝達の式は

$$\frac{d\mathcal{I}_\nu}{\kappa_\nu \rho ds} = -\mathcal{I}_\nu + \mathcal{J}_\nu \quad (6.20)$$

とかける。さらに、光学的深さ

$$d\tau = -\kappa_\nu \rho dz \quad (6.21)$$

を定義することで、ある軸 z をとって $ds = \mu dz$ と $\mu = \cos \theta$ を用いて

$$\mu \frac{d\mathcal{I}_\nu}{d\tau} = \mathcal{I}_\nu - \mathcal{J}_\nu \quad (6.22)$$

となる。これを Schwarzschild equation という。

電磁波の減光 (extinction) には、光が吸収 (absorption) され真に消失する効果と、異なる方向へと散乱 (scattering) されて消失する効果の二種類の和

$$\text{extinction} = \text{absorption} + \text{scattering} \quad (6.23)$$

となる。光子が原子・分子を電離し、原子の電離エネルギーと電離した電子の運動エネルギーとなる場合や、光子により原子・分子の電子が励起され、これが原子や分子同士の衝突により脱励起されることにより熱化する場合などは吸収に対応する。散乱は、光子が原子・分子の電子を励起した後、そのまま脱励起し同じ周波数の光子を出す場合や、電子、原子・分子による散乱などが含まれる。opacity の吸収、散乱による成分を、それぞれ、true absorption coefficient μ_a と scattering coefficient μ_s で表す。

$$\kappa_\nu = \mu_a + \mu_s \quad (6.24)$$

となる。このとき、emission coefficient は mean intensity

$$J_\nu \equiv \langle \mathcal{I}_\nu \rangle = \frac{1}{4\pi} \int P(\Omega) d\Omega \mathcal{I}_\nu \quad (6.25)$$

を用いて ($P(\Omega)$ は散乱の特性関数)

$$\eta_\nu = \mu_a B_\nu + \mu_s J_\nu \quad (6.26)$$

と書ける。つまり放射源関数は

$$\mathcal{J}_\nu = \frac{\mu_a B_\nu + \mu_s J_\nu}{\mu_a + \mu_s} = (1 - \omega_0) B_\nu + \omega_0 J_\nu \quad (6.27)$$

と書ける。ここに

$$\omega_0 \equiv \frac{\mu_s}{\mu_a + \mu_s} \quad (6.28)$$

は単散乱アルベドと呼ばれる。つまり散乱のある場合の放射伝達式は、

$$\mathcal{J}_\nu = \omega_0 J_\nu + (1 - \omega_0) B_\nu \quad (6.29)$$

となる。

6.3 輻射光スペクトルのモデリング

輻射光スペクトルの生成は放射伝達を解く必要がある。しかし散乱を無視できる場合、比較的簡単な放射伝達計算となる。散乱無し (pure absorption) の場合、Schwarzschild 方程式が形式的に積分可能となることから Intensity に関してレイヤー間を伝達させることができる。この場合、射出フラックスは、ToA での Intensity を角度方向に積分することにより最後に計算する。

大気上端から取った鉛直方向の optical depth を τ とする。

Schwarzschild Equation は

$$\mu \frac{dI_\nu}{d\tau} = I_\nu - \mathcal{J}_\nu \quad (6.30)$$

であるから、 $\tau' = \tau/\mu$ と定義することで、 I_ν を μ の関数として

$$\dot{I}_\nu(\mu) = I_\nu(\mu) - \mathcal{J}_\nu(\mu) \quad (6.31)$$

とおける。両辺に $e^{-\tau'}$ をかけることで、 $\tau' = \tau'_A$ から $\tau' = \tau'_B$ まで積分を実行することができ、形式解

$$I_\nu(\tau'_B; \mu) e^{-\tau'_B} = I_\nu(\tau'_A, \mu) e^{-\tau'_A} - \int_{\tau'_A}^{\tau'_B} \mathcal{J}_\nu(\mu) e^{-\tau'} d\tau' \quad (6.32)$$

が得られる。これは積分系の Schwarzschild equation である。波数の添え字を省略する。また散乱無しの放射を考える ($\omega_0 = 0$) ので、式 (6.29) より、 $\mathcal{J}_\nu(\mu)$ は角度依存性のないプランク関数によって置き換えられる。

outgoing に寄与する intensity のみ、つまり $\mu \geq 0$ のみを考えることにして $I(\tau, \mu) = I^+(\tau, \mu)$ という表

記を用いる^{*3}。ここである地点で optical depth の基準をとり $\tau_B = 0$ として

$$I^+(0; \mu) = I^+(\tau'_A, \mu) e^{-\tau'_A} + \int_0^{\tau'_A} B_\nu(\tau) e^{-\tau'} d\tau' \quad (6.33)$$

が得られる。

式 (6.33) にて $\tau_A = \tau_s$ (下端) として、ある波数での放射伝達を Transmission を $T(z; \mu) \equiv e^{-\tau(z)/\mu}$ として ($z = 0$ が大気下端)、以下の量から計算することができる。

$$I^+(\mu) = B_\nu(T_s) T(z=0; \mu) + \int_0^\infty B_\nu(T) \frac{dT(z)}{dz} dz \quad (6.34)$$

T_s は大気下端温度、 $\tau'_s = \tau_s/\mu$ と定義した。上の式を $T(z)$ に関して離散化して

$$I^+(\mu) = B_\nu(T_s) T_{N_{\text{layer}}-1}(\mu) + \sum_{j=0}^{N_{\text{layer}}-1} B_\nu(T_j) (T_{j-1}(\mu) - T_j(\mu)) \quad (6.35)$$

のように計算している^{*4}。透過関数 $T_j = e^{\sum_{i=0}^j -\tau(z)/\mu}$ であることからわかるように、式 (6.35) の評価のために事前に透過関数を $0-j$ まで計算しておく必要がある。

これを μ について複数の求積点をもちいて計算する。

$$F^+(0) = 2\pi \int_0^1 \mu I^+(\mu) d\mu \quad (6.36)$$

これは N 流の産卵なしの Intensity に戻づく輻射計算に対応し、散乱系への拡張はそのままでは困難である。

6.4 散乱・反射スペクトルのモデリング (二流近似)[‡]

さて散乱・反射がある場合、flux に基づく二流近似がよく使われる。ここでは

^{*3} この意味では $\mu < 0$ に対しては $I(\tau, \mu) = -I^-(\tau, \mu)$ と定義される。

^{*4} Irwin+ では、大気下端から昇順にレイヤーインデックスを付与しているが、本文章では上端から昇順にインデックスを付与する。

平行平板の放射伝達の式は

$$\mu \frac{d\mathcal{I}_\nu}{d\tau} = \mathcal{I}_\nu - \mathcal{J}_\nu \quad (6.37)$$

であった。

まず extinction としては、吸収と散乱を両方考えることから

$$\kappa_\nu = \mu_a + \mu_s \quad (6.38)$$

とおける。ここに true absorption coefficient μ_a 、(scattering coefficient μ_s を用いている。

放射としては熱放射 C と下からの散乱 D を含めたものを考えよう。E から H までは外からの光 (外部光) を考えるケースであり、恒星の散乱光を考える場合に必要であるが今は考えないこの場合、emission coefficient は放射によるものと散乱によるものの和であるので

$$\eta_\nu = \mu_a B_\nu + \mu_s \frac{1}{4\pi} \int d\Omega \mathcal{P}(\Omega) \mathcal{I}_\nu \quad (6.39)$$

と書ける。右辺第一項では、第2項の散乱を無視した場合、吸収 = 放射 (キルヒホッフの法則) となっている。すなわち局所熱力学平衡がなりたっているという仮定がおかれている^{*5}。そのため、右辺第一項にプランク関数が現れている。局所熱力学平衡と詳細つり合いから導かれる強度分布がプランク分布だからである。また、第2項は散乱を表す項であり、 $\mathcal{P}(\Omega)$ は散乱の方向依存性をあらわす関数である。ここでは例えば薄い雲による散乱を想定しておこう。

放射源関数は

$$\mathcal{J}_\nu = \frac{\eta_\nu}{\kappa_\nu} = (1 - \omega_0) B_\nu + \frac{\omega_0}{4\pi} \int d\Omega \mathcal{P}(\Omega) \mathcal{I}_\nu \quad (6.40)$$

ここに

$$\omega_0 \equiv \frac{\mu_s}{\mu_a + \mu_s} \quad (6.41)$$

は単散乱アルベドである。

モーメントを用いた記法、すなわち μ の n 乗をかけて立体角積分をした量 (モーメント) を用いて、specific intensity の立体角方向の分布を表現する方

^{*5} 大気上層、例えば地球での中間圏あたりより上層ではこの仮定は成り立たない

法での記述が可能となる。全球面上での積分平均を定義しておこう：

$$\langle \mathcal{F} \rangle \equiv \frac{1}{4\pi} \int d\Omega \mathcal{P}(\Omega) \mathcal{F} = \frac{1}{4\pi} \int d\phi \int_{-1}^1 d\mu \mathcal{P}(\Omega) \mathcal{F} \quad (6.42)$$

とする。0,1,2 次のモーメントを次の記号で略記する。

$$J_\nu \equiv \langle \mathcal{I}_\nu \rangle \quad (6.43)$$

$$H_\nu \equiv \langle \mu \mathcal{I}_\nu \rangle \quad (6.44)$$

$$K_\nu \equiv \langle \mu^2 \mathcal{I}_\nu \rangle \quad (6.45)$$

0 次のモーメントを用いて、放射源関数は

$$\mathcal{J}_\nu = \frac{\eta_\nu}{\kappa_\nu} = \omega_0 J_\nu + (1 - \omega_0) B_\nu \quad (6.46)$$

と書ける。

放射伝達の式は

$$\mu \frac{d\mathcal{I}_\nu(\Omega)}{d\tau} = \mathcal{I}_\nu(\Omega) - \omega_0 J_\nu - (1 - \omega_0) B_\nu \quad (6.47)$$

となる。

二流近似では、これを上流側のフラックスと下流側に分けることを考える。上半球 (US) と下半球 (LS) で積分する。今、モーメント方程式を意識して、任意の関数 \mathcal{F} に対する積分平均演算子を

$$\begin{aligned} \langle \mathcal{F} \rangle_{\text{US}} &\equiv \frac{1}{4\pi} \int_{\text{US}} d\Omega \mathcal{P}(\Omega) \mathcal{F} \\ &= \frac{1}{4\pi} \int d\phi \int_0^1 d\mu \mathcal{P}(\Omega) \mathcal{F} \end{aligned} \quad (6.48)$$

$$\begin{aligned} \langle \mathcal{F} \rangle_{\text{LS}} &\equiv -\frac{1}{4\pi} \int_{\text{LS}} d\Omega \mathcal{P}(\Omega) \mathcal{F} \\ &= \frac{1}{4\pi} \int d\phi \int_0^{-1} d\mu \mathcal{P}(\Omega) \mathcal{F} \end{aligned} \quad (6.49)$$

と表記する。二流近似では下向きの量も正值になるように、後者には負号をつけて定義してあること、つまり

$$\langle \mathcal{F} \rangle = \langle \mathcal{F} \rangle_{\text{US}} - \langle \mathcal{F} \rangle_{\text{LS}} \quad (6.50)$$

に注意する。さて大気にたいし上向きに射出するフラックスは、強度に対し上方向の角度依存性をかけ上半球について積分したものであるから、すなわち

$$F^+ = \int_{\text{US}} d\Omega \mu \mathcal{P}(\Omega) \mathcal{I}_\nu(\Omega) = 4\pi \langle \mu \mathcal{I}_\nu(\Omega) \rangle_{\text{US}} \quad (6.51)$$

となる。 $\mathcal{I}_\nu(\Omega) = \mathcal{I}_\nu(\mu)$ と書ける。

いま等方散乱 $\mathcal{P}(\Omega) = 1$ かつ強度の azimuth 依存性が無視できるなら、

$$F^+ = 2\pi \int_0^1 d\mu \mu \mathcal{I}_\nu(\mu) \quad (6.52)$$

となる。

散乱無しの直接解と透過関数

式 (6.47) は散乱無し ($\omega_0 = 0$) の場合、立体角の azimuth 方向を無視し、 μ だけに依存するとする ($\mathcal{I}_\nu(\Omega) = \mathcal{I}_\nu(\mu)$) と、両辺に $e^{-\tau/\mu}$ を掛けて積分することにより $\mathcal{I}_\nu(\mu)$ について解析的に解ける。すなわち

$$\frac{d}{d\tau} \left(\mathcal{I}_\nu(\tau, \mu) e^{-\tau/\mu} \right) = -\frac{B_\nu(T(\tau))}{\mu} e^{-\tau/\mu} \quad (6.53)$$

を τ で積分すればよい。ここで図6.3のような、一層の大気層を考え、層の下部の intensity を $\mathcal{I}_\nu(\tau_1, \mu) = \mathcal{I}_1(\mu)$ 、上部を $\mathcal{I}_\nu(\tau_2, \mu) = \mathcal{I}_2(\mu)$ とおく。式 (6.53) より

$$\mathcal{I}_1(\mu) = \mathcal{I}_2(\mu) e^{-\Delta\tau/\mu} + \frac{1}{\mu} \int_{\tau_1}^{\tau_2} d\tau B_\nu(T(\tau)) e^{-(\tau-\tau_1)/\mu} \quad (6.54)$$

となる。ここに $\Delta\tau = \tau_2 - \tau_1$ である。層上部からの上向きフラックス $F_{+,1}$ を、 $2\pi\mu$ をかけて上半球で積分すると

$$F_{+,1} = 2\pi \int_0^1 d\mu \mu \mathcal{I}_1(\mu) \quad (6.55)$$

$$\begin{aligned} &= 2\pi \int_0^1 d\mu \mathcal{I}_2(\mu) \mu e^{-\Delta\tau/\mu} \\ &\quad + \int_{\tau_1}^{\tau_2} d\tau \pi B_\nu(T(\tau)) \mathcal{G}(\tau - \tau_1) \end{aligned} \quad (6.56)$$

$$\mathcal{G}(\tau) \equiv 2 \int_0^1 d\mu e^{-\tau/\mu} \quad (6.57)$$

となる。この式は割と見通しが良い。すなわち上部から射出されるフラックスは、下部からの上向きの強度の伝達分 ($\Delta\tau/\mu$ の減光を受けたものに対し μ をかけ上半球積分したもの) と層内からの黒体輻射に対し順次、 $\mathcal{G}(x)$ (図6.4参照) の減光がかかって上部まで達したものの和となっている。特に大気上端を上部とし、すなわち $\tau_1 = 0$ ととり、下端からの放射を無視す

る、すなわち $\mathcal{I}_{\nu 2} = 0$ とした場合、式 (6.59) は

$$F_{+,1} = \int_0^\infty d\tau \pi B_\nu(T(\tau)) \mathcal{G}(\tau) \quad (6.58)$$

となる。

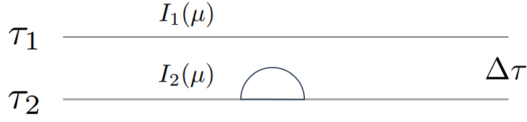


図6.3

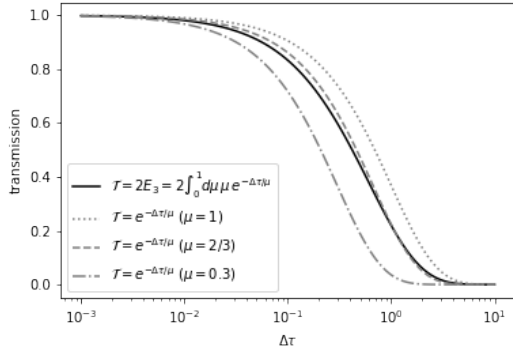


図6.4 吸収のみの場合の透過関数。

散乱ありの場合との対応をとるために、式 (6.59) を F^+ の漸化式の形で書いてみよう。このために 1-2 間を大気中の薄い層とみなす。この層の中では温度一定とする。また下からの強度 $\mathcal{I}_2(\mu)$ の μ 依存性を無視すると

$$F_{+,1} = 2\pi \mathcal{I}_2 \int_0^1 d\mu \mu e^{-\Delta\tau/\mu} + \pi B_\nu(T)() \quad (6.59)$$

$$(6.60)$$

Moment Closure による近似

(e.g. Meadow & Weaver 80 [11], Toon et al. 1989 [13] などを参照)。まず上半球 (US) と下半球 (LS) に対し、0 次のモーメント (平均強度)

$$J_{\nu+} \equiv \langle \mathcal{I}_\nu \rangle_{\text{US}} \quad (6.61)$$

$$J_{\nu-} \equiv \langle \mathcal{I}_\nu \rangle_{\text{LS}} \quad (6.62)$$

、一次のモーメント、

$$H_{\nu+} \equiv \langle \mu \mathcal{I}_\nu \rangle_{\text{US}} \quad (6.63)$$

$$H_{\nu-} \equiv \langle \mu \mathcal{I}_\nu \rangle_{\text{LS}} \quad (6.64)$$

。また上向き下向きのフラックスは

$$F^+ = 4\pi H_+ \quad (6.65)$$

$$F_- = 4\pi H_- \quad (6.66)$$

となる。

放射伝達の式は

$$\begin{aligned} \frac{d}{d\tau} \langle \mu \mathcal{I}_\nu(\Omega) \rangle_{\text{US}} &= \langle \mathcal{I}_\nu(\Omega) \rangle_{\text{US}} - \omega_0 \langle \mu \rangle_{\text{US}} J_\nu - (1 - \omega_0) \langle \mu \rangle_{\text{US}} B_\nu \end{aligned} \quad (6.67)$$

$$\begin{aligned} \frac{d}{d\tau} \langle \mu^2 \mathcal{I}_\nu(\Omega) \rangle_{\text{LS}} &= \langle \mu \mathcal{I}_\nu(\Omega) \rangle_{\text{LS}} - \omega_0 \langle \mu \rangle_{\text{LS}} J_\nu - (1 - \omega_0) \langle \mu \rangle_{\text{LS}} B_\nu \end{aligned} \quad (6.68)$$

となる。ここで $\langle \mu \rangle_{\text{US}} = -\langle \mu \rangle_{\text{LS}} = 1/2$ であるから、

$$\frac{d}{d\tau} \langle \mu \mathcal{I}_\nu(\Omega) \rangle_{\text{US}} = \langle \mathcal{I}_\nu(\Omega) \rangle_{\text{US}} - \frac{\omega_0}{2} J_\nu - \frac{(1 - \omega_0)}{2} B_\nu \quad (6.69)$$

$$\frac{d}{d\tau} \langle \mu \mathcal{I}_\nu(\Omega) \rangle_{\text{LS}} = \langle \mathcal{I}_\nu(\Omega) \rangle_{\text{LS}} + \frac{\omega_0}{2} J_\nu - \frac{(1 - \omega_0)}{2} B_\nu \quad (6.70)$$

ここで $J_\nu = J_{\nu+} - J_{\nu-}$ を使い、また closure relation を

$$\eta_+ = J_{\nu+}/H_{\nu+} \quad (6.71)$$

$$\eta_- = -J_{\nu-}/H_{\nu-} \quad (6.72)$$

と置くことで、

$$\dot{H}_{\nu+} = \eta_+ \left(1 - \frac{\omega_0}{2}\right) H_{\nu+} - \frac{\omega_0}{2} H_{\nu-} - \frac{(1 - \omega_0)}{2} B_\nu \quad (6.73)$$

$$\dot{H}_{\nu-} = \frac{\eta_+ \omega_0}{2} H_{\nu+} - \eta_- \left(1 - \frac{\omega_0}{2}\right) H_{\nu-} + \frac{(1 - \omega_0)}{2} B_\nu \quad (6.74)$$

または

$$\dot{F}^+ = \eta_+ \left(1 - \frac{\omega_0}{2}\right) F^+ - \frac{\eta_- \omega_0}{2} F_- - 2\pi(1 - \omega_0) B_\nu \quad (6.75)$$

$$\dot{F}^- = \frac{\eta_+ \omega_0}{2} F^+ - \eta_- \left(1 - \frac{\omega_0}{2}\right) F_- + 2\pi(1 - \omega_0) B_\nu \quad (6.76)$$

となる。

ここで上半球と下半球の closure relation が同じ形とおもうと $\eta_+ = \eta_- \equiv \eta$ となり

$$\dot{F}_+(\tau) = \gamma_1 F^+(\tau) - \gamma_2 F_-(\tau) - 2\pi(1 - \omega_0)B_\nu(\tau) \quad (6.77)$$

$$\dot{F}_-(\tau) = \gamma_2 F^+(\tau) - \gamma_1 F_-(\tau) + 2\pi(1 - \omega_0)B_\nu(\tau) \quad (6.78)$$

の形となる。ただし

$$\gamma_1 \equiv \eta \left(1 - \frac{\omega_0}{2}\right) \quad (6.79)$$

$$\gamma_2 \equiv \frac{\eta \omega_0}{2} \quad (6.80)$$

とおいた。 $\gamma_1 + \gamma_2 = \eta$, $\gamma_1 - \gamma_2 = \eta(1 - \omega_0)$ である。

式 (6.77, 6.78) は一階の連立微分方程式であり、齊次系に $B_\nu(\tau)$ 由来の τ 依存性のある項が付加されたものである。よって $B_\nu(\tau)$ をテイラー展開して打ち切り有限項の多項式に近似すれば解くことができる。

$\gamma_1 = 2 - \omega_0$, $\gamma_2 = \omega_0$ ととれば、半球平均近似 (Hemispheric mean) かつ非対称パラメタ $g = 0$ の時と一致する [13]。式 (6.77)、(6.78) を解くために、新たなパラメタ $F_{\text{sum}} \equiv F^+ + F_-$, $F_{\text{net}} \equiv F^+ - F_-$ と定義し書き換えると以下を得る：

$$\dot{F}_{\text{net}} = (\gamma_1 - \gamma_2)F_{\text{sum}} - 4\pi(1 - \omega_0)B_\nu \quad (6.81)$$

$$\dot{F}_{\text{sum}} = (\gamma_1 + \gamma_2)F_{\text{net}}, \quad (6.82)$$

これより二階微分方程式

$$\ddot{F}_{\text{sum}} - (\gamma_1^2 - \gamma_2^2)F_{\text{sum}} + 4\pi(1 - \omega_0)(\gamma_1 + \gamma_2)B_\nu = 0. \quad (6.83)$$

が得られる。

ここで B_ν の変数 τ に対するテイラー展開の一次まで残すことを考えよう [13] [4]。

$$B_\nu \approx B_0 + \dot{B}(\tau - \tau_0). \quad (6.84)$$

この時、解は以下のように求まる。

$$F_{\text{sum}} = c_1 e^{\lambda\tau} + c_2 e^{-\lambda\tau} + \frac{4\pi(1 - \omega_0)}{\gamma_1 - \gamma_2} B_\nu \quad (6.85)$$

$$F_{\text{net}} = c_1 \delta e^{\lambda\tau} - c_2 \delta e^{-\lambda\tau} + \frac{4\pi(1 - \omega_0)}{\gamma_1^2 - \gamma_2^2} \dot{B}_\nu, \quad (6.86)$$

ここに

$$\lambda \equiv \sqrt{\gamma_1^2 - \gamma_2^2} \quad (6.87)$$

$$\delta \equiv \sqrt{\frac{\gamma_1 - \gamma_2}{\gamma_1 + \gamma_2}}. \quad (6.88)$$

上の解を F^+ , F_- に戻すことで以下が求まる。

$$F^+(\tau) = c_1 \zeta_+ e^{\lambda\tau} + c_2 \zeta_- e^{-\lambda\tau} + \pi \mathcal{B}_+(\tau) \quad (6.89)$$

$$F^-(\tau) = c_1 \zeta_- e^{\lambda\tau} + c_2 \zeta_+ e^{-\lambda\tau} + \pi \mathcal{B}_-(\tau) \quad (6.90)$$

ここに

$$\mathcal{B}_\pm(\tau) \equiv \frac{2(1 - \omega_0)}{\gamma_1 - \gamma_2} \left(B_\nu(\tau) \pm \frac{1}{\gamma_1 + \gamma_2} \dot{B}_\nu(\tau) \right) \quad (6.91)$$

$$\zeta_\pm \equiv \frac{1}{2}(1 \pm \delta). \quad (6.92)$$

である。 ζ_\pm は結合定数と呼ばれる [4]。

大気レイヤーモデルの放射伝達の解法

ここまでは Hemispheric mean の二流近似を導出してきたが、Toon89 型の二流近似における上向きおよび下向きストリームの解は以下のように同様に表現できる。

$$F^+(\tau) = c_1 \zeta^+ e^{\lambda\tau} + c_2 \zeta^- e^{-\lambda\tau} + \pi \mathcal{B}^+(\tau) \quad (6.93)$$

$$F^-(\tau) = c_1 \zeta^- e^{\lambda\tau} + c_2 \zeta^+ e^{-\lambda\tau} + \pi \mathcal{B}^-(\tau) \quad (6.94)$$

ここで、 ζ^\pm は結合係数と呼ばれる。注目すべきことに、球面調和関数 (SH) 法の二流近似も同じ形をとる。そこでこの形の方程式から大気レイヤーモデルの放射伝達の解法を考えよう。

上記の方程式において、 ζ^\pm と λ は、 F^\pm の微分方程式の係数である γ_1 と γ_2 以下のように関連している。

$$\zeta^\pm = \frac{1}{2} \left(1 \pm \sqrt{\frac{\gamma_1 - \gamma_2}{\gamma_1 + \gamma_2}} \right) \quad (6.95)$$

$$\lambda = \sqrt{\gamma_1^2 - \gamma_2^2}, \quad (6.96)$$

係数 γ_1 と γ_2 は、単一散乱アルベド ω と非対称パラメータ g によって決定される。この関係はモーメント閉包の方法に依存する [13]。半球平均を使用する場合、 $\gamma_1 = 2 - \omega(1 + g)$ および $\gamma_2 = \omega(1 - g)$ の関係

が成り立つ。また、簡約化された源関数は以下のように表される

$$\mathcal{B}^\pm(\tau) = \frac{2(1-\omega)}{\gamma_1 - \gamma_2} \left(B_\nu(\tau) \pm \frac{1}{\gamma_1 + \gamma_2} \dot{B}_\nu(\tau) \right). \quad (6.97)$$

上記の方程式において、第二項は等温層の場合には無視される。

式 (6.93, 6.94) は以下のように表現できる

$$\mathbf{F}(\tau) = Q(\tau)\mathbf{x} + \pi\mathcal{B}(\tau) \quad (6.98)$$

ここで $\mathbf{F}(\tau) = (F^+(\tau), F^-(\tau))^\top$, $\mathcal{B}(\tau) = (\mathcal{B}^+(\tau), \mathcal{B}^-(\tau))^\top$, $\mathbf{x} \equiv (c_1, c_2)^\top$, そして

$$Q(\tau) = \begin{pmatrix} \zeta^+ e^{\lambda\tau} & \zeta^- e^{-\lambda\tau} \\ \zeta^- e^{\lambda\tau} & \zeta^+ e^{-\lambda\tau} \end{pmatrix} \quad (6.99)$$

次に、光学的厚さの差 $\Delta\tau_n$ を持つ N 層モデルを考え、第 n 層の上端における光学的深度を $\tau = \tau_n = \sum_{i=0}^{n-1} \Delta\tau_i$ ($n \geq 1$) および $\tau_0 = 0$ として定義する。

第 n 層について考えよう。

$$\mathbf{F}(\tau_n) = Q_n(\tau_n)\mathbf{x}_n + \pi\mathcal{B}_n(\tau_n) = \mathbf{F}_n \quad (6.100)$$

$$\begin{aligned} \mathbf{F}(\tau_n + \Delta\tau_n) &= Q_n(\tau_n + \Delta\tau_n)\mathbf{x}_n + \pi\mathcal{B}_n(\tau_n + \Delta\tau_n) \\ &= \mathbf{F}_{n+1} \end{aligned} \quad (6.101)$$

ここで Q_n は第 n 層の $Q(\tau)$ として定義される。したがって、第 n 層の内部パラメータ ζ_n^\pm と λ_n を設定すると、

$$Q_n(\tau) = \begin{pmatrix} \zeta_n^+ e^{\lambda_n \tau} & \zeta_n^- e^{-\lambda_n \tau} \\ \zeta_n^- e^{\lambda_n \tau} & \zeta_n^+ e^{-\lambda_n \tau} \end{pmatrix}. \quad (6.102)$$

式 (6.100) と (6.101) から、漸化関係の線形形式を得る：

$$\mathbf{F}_{n+1} = \mathcal{G}(\Delta\tau_n)\mathbf{F}_n + \pi\mathbf{G}_n, \quad (6.103)$$

ここで

$$\mathbf{G}_n \equiv \mathcal{B}_n(\tau_n + \Delta\tau_n) - \mathcal{G}(\Delta\tau_n)\mathcal{B}_n(\tau_n) \quad (6.104)$$

かつ $\mathbf{G}_n = (G_n^+, G_n^-)^\top$ である。また、伝達関数を以下のように定義した

$$\mathcal{G}(\Delta\tau_n) \equiv Q(\tau_n + \Delta\tau_n)Q^{-1}(\tau_n) \quad (6.105)$$

$$= Q_n(\tau_n) \begin{pmatrix} e^{\lambda_n \Delta\tau_n} & 0 \\ 0 & e^{-\lambda_n \Delta\tau_n} \end{pmatrix} Q_n^{-1}(\tau_n) \quad (6.106)$$

これは $\mathcal{G}(\Delta\tau_n)\mathbf{q}_i = \lambda_i' \mathbf{q}_i$ の固有値分解であり、ここで \mathbf{q}_i は $Q_n(\tau_n)$ の第 i 列ベクトルである。したがって \mathbf{q}_i を正規化でき、特に \mathbf{q}_1 を $e^{\lambda_n \tau_n}$ で、 \mathbf{q}_2 を $e^{-\lambda_n \tau_n}$ で正規化できる。以下のように定義することで

$$Z_n \equiv Q_n(0) = \begin{pmatrix} \zeta_n^+ & \zeta_n^- \\ \zeta_n^- & \zeta_n^+ \end{pmatrix}, \quad (6.107)$$

上記の方程式を以下のように書き直せる

$$\begin{aligned} \mathcal{G}(\Delta\tau_n) &= Z_n \begin{pmatrix} e^{\lambda_n \Delta\tau_n} & 0 \\ 0 & e^{-\lambda_n \Delta\tau_n} \end{pmatrix} Z_n^{-1} \\ &= \frac{1}{\zeta_n^{+2} - \zeta_n^{-2}} \\ &\quad \times \begin{pmatrix} \zeta_n^{+2} e^{t_n} - \zeta_n^{-2} e^{-t_n} & -\zeta_n^+ \zeta_n^- (e^{t_n} - e^{-t_n}) \\ \zeta_n^+ \zeta_n^- (e^{t_n} - e^{-t_n}) & \zeta_n^{+2} e^{-t_n} - \zeta_n^{-2} e^{t_n} \end{pmatrix}, \end{aligned} \quad (6.108)$$

ここで

$$t_n \equiv \lambda_n \Delta\tau_n \quad (6.110)$$

は $\Delta\tau_n$ の関数であるが、 τ_n の関数ではない。

等温層の場合、 $\mathcal{B}_n(\tau_n) = \mathcal{B}_n(\tau_n + \Delta\tau) \equiv \mathcal{B}_n \mathbf{u}$, $\mathcal{B}_n = \mathcal{B}^+(\tau_n) = \mathcal{B}^-(\tau_n)$ であり、ここで \mathbf{u} は単位ベクトル $\mathbf{u} \equiv (1, 1)^\top$ である。源行列 \mathbf{G}_n は以下のように簡約化できる

$$\mathbf{G}_n = \mathcal{B}_n(I - \mathcal{G}(\Delta\tau_n))\mathbf{u} \quad (6.111)$$

$$= \frac{\mathcal{B}_n}{\zeta_n^+ + \zeta_n^-} \begin{pmatrix} \zeta_n^+(1 - e^{t_n}) + \zeta_n^-(1 - e^{-t_n}) \\ \zeta_n^+(1 - e^{-t_n}) + \zeta_n^-(1 - e^{t_n}) \end{pmatrix} \quad (6.112)$$

ここで I は単位行列である。

単一層内での伝達

線形形式 (6.103) は数学的に簡潔であるが、その物理的意味は明確ではない。そこで、単一層内での光の伝達を表す形式に変換する。

式 (6.103) から以下が得られる

$$\mathbf{F}_n^+ = \mathcal{G}_{11}^{-1} \mathbf{F}_{n+1}^+ - \mathcal{G}_{11}^{-1} \mathcal{G}_{12} \mathbf{F}_n^- - \mathcal{G}_{11}^{-1} \pi \mathbf{G}_n^+, \quad (6.113)$$

ここで \mathcal{G}_{ij} は記号から n を省略した $\mathcal{G}(\Delta\tau_n)$ の $(i-j)$ 成分である。この方程式を $\mathbf{F}_{n+1}^- = \mathcal{G}_{21} \mathbf{F}_n^+ + \mathcal{G}_{22} \mathbf{F}_n^- + \pi \mathbf{G}_n^-$ に代入すると、以下を得る

$$\begin{aligned} \mathbf{F}_{n+1}^- &= \mathcal{G}_{21} \mathcal{G}_{11}^{-1} \mathbf{F}_{n+1}^+ + (\mathcal{G}_{22} - \mathcal{G}_{21} \mathcal{G}_{11}^{-1} \mathcal{G}_{12}) \mathbf{F}_n^- \\ &\quad + \pi \mathbf{G}_n^- - \mathcal{G}_{21} \mathcal{G}_{11}^{-1} \pi \mathbf{G}_n^+. \end{aligned} \quad (6.114)$$

二流の場合、式 (6.109) から $\mathcal{G}_{21} = -\mathcal{G}_{12}$ かつ $\mathcal{G}_{11}^{-1} = \mathcal{G}_{22} - \mathcal{G}_{21}\mathcal{G}_{11}^{-1}\mathcal{G}_{12} \equiv \mathcal{T}_n$ および $\mathcal{G}_{21}\mathcal{G}_{11}^{-1} = -\mathcal{G}_{11}^{-1}\mathcal{G}_{12} \equiv \mathcal{S}_n$ が導かれる。そこで、第 n 層内での放射伝達を以下のように表現できる

$$F_n^+ = \mathcal{T}_n F_{n+1}^+ + \mathcal{S}_n F_n^- - \mathcal{T}_n \pi G_n^+ \quad (6.115)$$

$$F_{n+1}^- = \mathcal{T}_n F_n^- + \mathcal{S}_n F_{n+1}^+ + \pi G_n^- - \mathcal{S}_n \pi G_n^+, \quad (6.116)$$

ここで

$$\mathcal{T}_n \equiv \frac{\zeta_n^{+2} - \zeta_n^{-2}}{\zeta_n^{+2} - (\zeta_n^- \mathcal{T}_n)^2} \mathcal{T}_n \quad (6.117)$$

$$\mathcal{S}_n \equiv \frac{\zeta_n^+ \zeta_n^-}{\zeta_n^{+2} - (\zeta_n^- \mathcal{T}_n)^2} (1 - \mathcal{T}_n^2) \quad (6.118)$$

は層の底部と頂部間の透過率およびフラックスの反対方向からの散乱とみなすことができる^{*6}。上記の式において、透過関数を定義した [4] :

$$\mathcal{T}_n \equiv e^{-\lambda_n \Delta \tau_n} \quad (6.122)$$

式 (6.115, 6.116) は本質的に [4] によって導出された二流近似の解析的表現と同じであることに注意されたい^{*7}。

等温層の場合、式 (6.115) と (6.116) を以下のように簡約化できる

$$F_n^+ = \mathcal{T}_n F_{n+1}^+ + \mathcal{S}_n F_n^- + \pi \hat{\mathcal{B}}_n \quad (6.123)$$

$$F_{n+1}^- = \mathcal{T}_n F_n^- + \mathcal{S}_n F_{n+1}^+ + \pi \hat{\mathcal{B}}_n \quad (6.124)$$

ここで

$$\hat{\mathcal{B}}_n \equiv (1 - \mathcal{T}_n - \mathcal{S}_n) \mathcal{B}_n. \quad (6.125)$$

^{*6} 式 (6.118) での散乱の不透明極限 ($\mathcal{T}_n = 0$) を $\mathcal{S}_\infty \equiv \zeta_- / \zeta_+$ として定義すると、式 (6.119) を [?] での flux-adding treatment で使用されるものと同様の形に書き直すことができる :

$$\mathcal{T}_n = \frac{\mathcal{S}_\infty (1 - e^{-2\lambda_n \Delta \tau_n})}{1 - \mathcal{S}_\infty^2 e^{-2\lambda_n \Delta \tau_n}}. \quad (6.119)$$

半球平均の場合、以下を得る

$$\mathcal{S}_\infty = \frac{\sqrt{1 - \omega g} - \sqrt{1 - \omega}}{\sqrt{1 - \omega g} + \sqrt{1 - \omega}} \quad (6.120)$$

$$\lambda_n = 2\sqrt{(1 - \omega g)(1 - \omega)}. \quad (6.121)$$

^{*7} [4] の式 (3.58)。我々の形式での λ_n は [4] での \mathcal{D} に対応する。

flux-adding treatment

散乱を含む二流近似の解法として、反射の形式を利用する flux-adding treatment [12, ?] が提案されている。これは古典的な加算技法から類推的に導出される。flux-adding treatment は、与えられた層における上向きフラックスが、下向きフラックスの反射と層からの源項の和として表現できると仮定する。

$$F_n^+ = \hat{R}_n^+ F_n^- + \hat{S}_n^+ \quad (6.126)$$

$$F_n^- = \hat{R}_n^- F_n^+ + \hat{S}_n^-. \quad (6.127)$$

式 (6.123) 中の F_n^+ を式 (6.126) で置き換え、 \mathcal{T}_n を乗じると、以下の式が導かれる。

$$\mathcal{T}_n^2 F_{n+1}^+ = (\hat{R}_n^+ - \mathcal{S}_n) \mathcal{T}_n F_n^- + \mathcal{T}_n (\hat{S}_n^+ - \pi \hat{\mathcal{B}}_n). \quad (6.128)$$

式 (6.124) を用いて前式から F_n^- を消去することにより、以下の結果が導かれる :

$$F_{n+1}^+ = \frac{\hat{R}_n^+ - \mathcal{S}_n}{\mathcal{T}_n^2 - \mathcal{S}_n^2 + \mathcal{S}_n \hat{R}_n^+} F_{n+1}^- + \frac{\hat{\mathcal{B}}_n (\mathcal{S}_n - \mathcal{T}_n - \hat{R}_n^+) + \mathcal{T}_n \hat{S}_n^+}{\mathcal{T}_n^2 - \mathcal{S}_n^2 + \mathcal{S}_n \hat{R}_n^+}. \quad (6.129)$$

前式の右辺第一項の係数を \hat{R}_{n+1}^+ 、第二項を \hat{S}_{n+1}^+ とする。これにより以下の漸化関係が導かれる。

$$\hat{R}_n^+ = \mathcal{S}_n + \frac{\mathcal{T}_n^2 \hat{R}_{n+1}^+}{1 - \mathcal{S}_n \hat{R}_{n+1}^+} \quad (6.130)$$

$$\hat{S}_n^+ = \hat{\mathcal{B}}_n + \frac{\mathcal{T}_n (\hat{S}_{n+1}^+ + \hat{\mathcal{B}}_n \hat{R}_{n+1}^+)}{1 - \mathcal{S}_n \hat{R}_{n+1}^+}. \quad (6.131)$$

式 (6.131) の導出において、式 (6.130) を用いて式 (6.129) の第二項中の \hat{R}_n^+ を消去した。

したがって、最下層 ($n = N - 1$) の底面境界の反射率 (すなわちアルベド) を仮定すれば、 \hat{R}_0^+ と \hat{S}_0^+ を計算でき、大気上端での出射フラックスを以下のように得ることができる

$$F_0^+ = \hat{R}_0^+ F_\star + \hat{S}_0^+, \quad (6.132)$$

ここで F_\star は入射恒星フラックスである。

出射フラックス計算には \hat{R}_n^- と \hat{S}_n^- を使用しないが、これらも導出できる。式 (6.124) において n を $n - 1$ で置き換え、次に式 (6.127) を用いて F_n^- を消

去する。さらに、式 (6.123) において n を $n-1$ で置換して $\mathcal{T}_{n-1}F_n^+$ を取り出し、これを上記の式に挿入すると、以下の式が導かれる。

$$F_{n-1}^- = \frac{\hat{R}_n^- - \mathcal{S}_{n-1}}{\mathcal{T}_{n-1}^2 - \mathcal{S}_{n-1}^2 + \mathcal{S}_{n-1}\hat{R}_n^-} F_{n-1}^+ + \frac{\hat{\mathcal{B}}_{n-1}(\mathcal{S}_{n-1} - \mathcal{T}_{n-1} - \hat{R}_n^-) + \mathcal{T}_{n-1}\hat{\mathcal{S}}_n^-}{\mathcal{T}_{n-1}^2 - \mathcal{S}_{n-1}^2 + \mathcal{S}_{n-1}\hat{R}_n^-} \quad (6.133)$$

右辺第一項の係数を \hat{R}_{n-1}^- 、第二項を $\hat{\mathcal{S}}_{n-1}^-$ とすることで、以下の漸化式が得られる。

$$\hat{R}_n^- = \mathcal{S}_{n-1} + \frac{\mathcal{T}_{n-1}^2 \hat{R}_{n-1}^-}{1 - \mathcal{S}_{n-1} \hat{R}_{n-1}^-} \quad (6.134)$$

$$\hat{\mathcal{S}}_n^- = \hat{\mathcal{B}}_{n-1} + \frac{\mathcal{T}_{n-1}(\hat{\mathcal{S}}_{n-1}^- + \hat{\mathcal{B}}_{n-1} \hat{R}_{n-1}^-)}{1 - \mathcal{S}_{n-1} \hat{R}_{n-1}^-} \quad (6.135)$$

式 (6.130)、(6.131)、(6.134)、(6.135) は本質的に [12] の式 (7)、(8)、(4)、(5) と同じである。

以上で見てきたように、flux-adding treatment では、散乱を各層の実効的な反射と源項の和として含めることにより、漸化関係を定式化し解くことができる。同様に、散乱を各層の実効的な透過と源項の和として含めることにより、他の漸化関係も定式化できる。この方法において実用上の利点をまだ特定していないが、フラックススペースの純粹吸収の自然な拡張であり理論的に興味深い。

参考文献

- [1] Mathieu Blondel and Vincent Roulet. The Elements of Differentiable Programming. *arXiv e-prints*, page arXiv:2403.14606, March 2024.
- [2] Nicholas N. Gibbons. equilibrium-c: A Lightweight Modern Equilibrium Chemistry Calculator for Hypersonic Flow Applications. *arXiv e-prints*, page arXiv:2412.07166, December 2024.
- [3] Sanford Gordon and Bonnie J McBride. Computer program for calculation of complex chemical equilibrium compositions and applications. part 1: Analysis. Technical report, 1994.
- [4] Kevin Heng. *Exoplanetary Atmospheres*. Princeton University Press, 2017.
- [5] Hajime Kawahara, Yui Kawashima, Shotaro Tada, Hiroyuki Tako Ishikawa, Ko Hosokawa, Yui Kasagi, Takayuki Kotani, Kento Masuda, Stevanus K. Nugroho, Motohide Tamura, Hibiki Yama, Daniel Kitzmann, Nicolas Minesi, and Brett M. Morris. Differentiable Modeling of Planet and Substellar Atmosphere: High-resolution Emission, Transmission, and Reflection Spectroscopy with ExoJAX2. *ApJ*, 985(2):263, June 2025.
- [6] Hajime Kawahara and Kento Masuda. Bayesian Dynamic Mapping of an Exo-Earth from Photometric Variability. *ApJ*, 900(1):48, September 2020.
- [7] Yui Kawashima, Hajime Kawahara, Yui Kasagi, Hiroyuki Tako Ishikawa, Kento Masuda, Takayuki Kotani, Tomoyuki Kudo, Teruyuki Hirano, Masayuki Kuzuhara, Stevanus K. Nugroho, John Livingston, Hiroki Harakawa, Jun Nishikawa, Masashi Omiya, Takuya Takarada, Motohide Tamura, and Akitoshi Ueda. Atmospheric Retrieval of Subaru/IRD High-resolution Spectrum of the Archetype T-type Brown Dwarf Gl 229 B. *ApJ*, 988(1):53, July 2025.
- [8] Kyriaki Kefala, Vincent Boudon, Sergei N. Yurchenko, and Jonathan Tennyson. Empirical rovibrational energy levels for methane. *JQSRT*, 316:108897, April 2024.
- [9] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *International Conference on Learning Representations (ICLR)*, 2015.
- [10] Anna Lueber and Adam J. Burgasser. Comparing Grid Model Fitting Methodologies for Low-temperature Atmospheres: Markov Chain Monte Carlo versus Random Forest Retrieval. *ApJ*, 988(1):31, July 2025.
- [11] W. E. Meador and W. R. Weaver. Two-stream approximations to radiative transfer in planetary atmospheres: a unified description of existing methods and a new improvement. *Journal of Atmospheric Sciences*, 37:630–643, March 1980.
- [12] Tyler D. Robinson and David Crisp. Linearized Flux Evolution (LiFE): A technique for rapidly adapting fluxes from full-physics radiative transfer models. *JQSRT*, 211:78–95, May 2018.
- [13] Owen B. Toon, C. P. McKay, T. P. Ackerman, and K. Santhanam. Rapid calculation of radiative heating rates and photodissociation rates in inhomogeneous multiple scattering atmospheres. *JGR*, 94:16287–16301, November 1989.

教科書類

ベイズ推定は Gelman らのベイズデータ解析 (BDA)、森北出版が定番である。また Kevin P. Murphy の Probabilistic Machine Learning はさらに網羅的に重要な事項がまとめられている。