

宇宙論と統計

1. はじめに

現在の観測的宇宙論は、非常にエキサイティングな研究分野です。宇宙背景放射（CMB）の宇宙論に代表されるように、高精度の測定結果から、ダークマター、ダークエネルギーの存在量、宇宙年齢など、人類の根源的な疑問に答えることに成功しています。21世紀の観測的宇宙論は、もはや理論と実験の垣根がない、言わば実証的科学とも言える分野に成長しています。

過去のIPMUニュースの記事でも、観測的宇宙論の進展を扱ったものがありますので、今回は少し変わった話をしたいと思います。それは宇宙論における統計的側面についてです。宇宙論の理論や解析には、統計的概念が多く取り入れられています。宇宙論の観測データは、一見複雑です。例えば、CMBの場合には、空の各方向からやってくるCMB光子の黒体輻射の温度が基本観測量になります。¹ 銀河サーベイの場合は、多数の銀河の空間分布が基本観測量です。これら宇宙論データをどのようにして定量化し、理論モデルと比較できるのでしょうか？ この宇宙論解析で用いられる作業過程、またその前提になっている仮定、さらにはその限界は何でしょうか？ これらの疑問を解説することが今回の記事の主題です。ここで解説すること以外にも、宇宙論と統計の境界領域の進展は近年目覚ましいものがあります。この記事で触れるテーマは、そのごく一部であることに注意して、話を進めたいと思います。

2. 宇宙原理、エルゴード仮定と 2 点相関関数

まず、議論のために、 $F(\theta)$ を宇宙論データから得られる「ゆらぎ場」としましょう。例えば、CMBの場合には、温度ゆらぎ場に対応し、 $F(\theta) = [T(\theta) - \bar{T}] / \bar{T}$ です。ここで $T(\theta)$ は空の θ 方向のCMB温度、 \bar{T} はCMBの平均温度（全天空平均した温度）です（図1の左上図参照）。3次元空間のゆらぎ場についても同様に定義できますが、とりあえずは2次元の場合を考えましょう。² 実際の観測では、ゆらぎ場 $F(\theta)$ は有限の角度分解能で測定されますので、 $F(\theta)$ は離散的なピクセル形式で与えられているとします。最新のCMB衛星Planckでは、約5分角³の角度分解能で、約40,000平方度の全天をサーベイしましたので、CMB温度ゆらぎ場は約500万ピクセルのデータで与えられ、膨大なデータ量になります。

しかし、宇宙論の理論モデルは観測した $F(\theta)$ 場を忠実に再現することはできません。より正確に言えば、再現するモデルを構築するためには、データの自由度に匹敵するパラメータを導入する必要があり、見返りの小さい、無駄な努力を払うことになります。そこで、宇宙論の解析では、**宇宙原理**を仮定するのが通常です。

- 宇宙は統計平均的な意味で一様かつ等方である

¹ 今回の記事では、主に宇宙論のゆらぎ場に注目します。

² 以下の議論では、空間の次元数は本質的ではありませんので、3次元の場合についても同様の結論が得られます。

³ 1分角は1度角の1/60の大きさ。

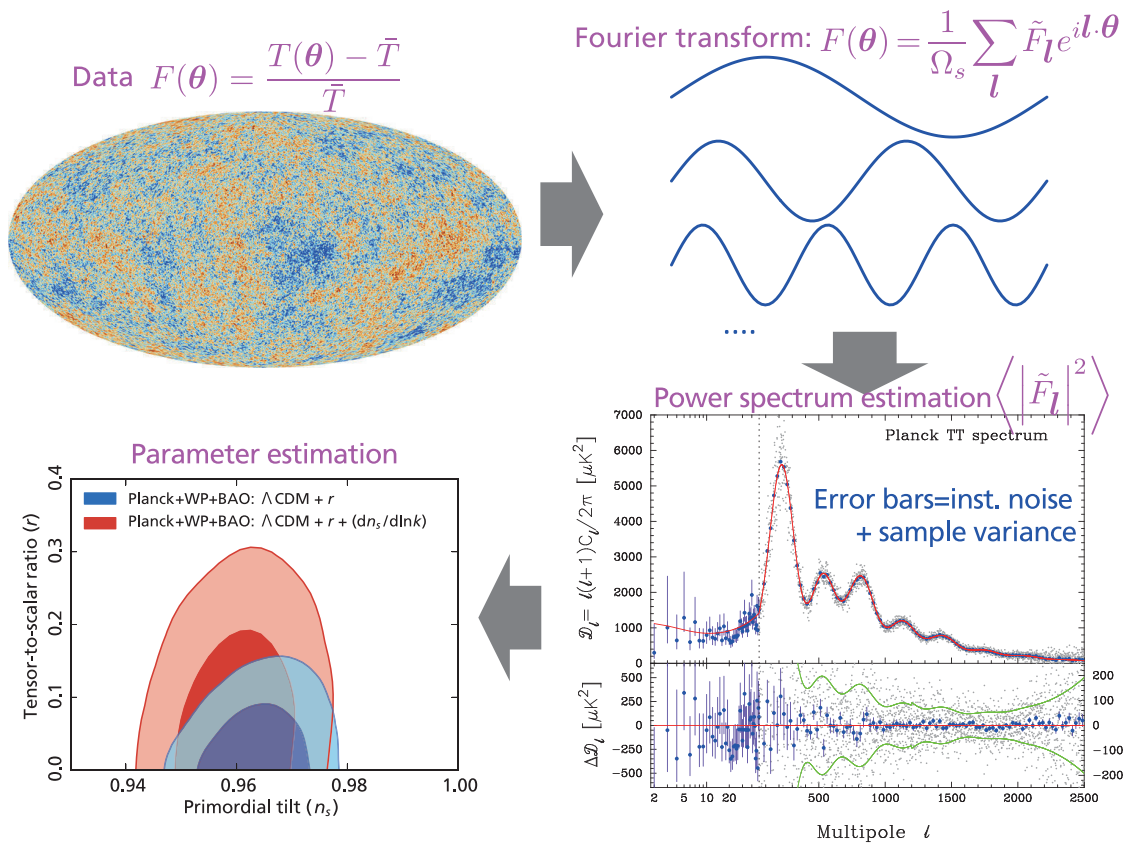


図1 宇宙論解析の流れ (2次元データのCMBの例)。宇宙論ゆらぎ場 $F(\theta)$ の測定 (左上図)。ゆらぎ場のフーリエ波数分解 (右上図)。右下図は、パワースペクトル推定の結果。灰色点は各波数ビンでの測定値。エラーバーを伴う青点は、複数の波数ビンで測定値を平均した結果。エラーバーとして、有限数のフーリエモードしか測定できないという事実から生じる標本分散および検出器などに付随するエラーが考慮されている。実線はベストフィットの Λ CDMモデルの理論予言。統計誤差内で観測データを良く再現している。左下図は、測定とモデルの比較により得られた宇宙論パラメータの推定結果の例。パワースペクトルの統計誤差をパラメータの決定の信頼区間に伝播させている。

という原理です。要するに、宇宙には特別な場所あるいは方向が存在しない、あるいは我々の天の川銀河は宇宙のなかで特別な位置に存在するわけではない、という民主的な考えです。

この宇宙原理を受け入れれば、我々が観測した $F(\theta)$ 場は、広大な宇宙のなかに存在する母集団の $\{F(\theta)\}$ から得られた典型的な標本サンプルと考えられます。CMBの例で言えば、宇宙のなかで遠く離れたところにいる観測者 (人類とは限りませんが) がCMBを観測したとしても、我々が見ているCMB温度ゆらぎ場と「大体」同じものを見るだろうと考えるわけです。この「同じ」程度、つまりどのくらい我々の観測したCMB場が「典型的」かを確率的に定量化する必要があります。

観測したゆらぎ場を定量化するために、フーリエ分

解を考えましょう。フーリエ分解とは、 $F(\theta)$ 場を以下のようにモード分解する方法です (図1の右上図)。

$$F(\theta) = \frac{1}{\Omega_s} \sum_l \tilde{F}_l e^{i.l.\theta} \quad (1)$$

簡単のため、天球の曲率は無視し、天球上の観測領域を2次元平面と近似することにします。観測領域の天球上での面積を Ω_s とすれば、フーリエ分解の最小波数、いわゆる基本波数は $l_f \approx 2\pi/\Omega_s^{1/2}$ ($\Omega_s^{1/2}$ は面積の一辺の長さ) となります。この基本波数の整数倍の波数、つまり $l = l_f (n_x, n_y)$ ($n_x, n_y = \pm 1, \pm 2, \dots$) の波数で、ゆらぎ場 $F(\theta)$ をモード分解します。フーリエ係数 \tilde{F}_l は、 $F(\theta)$ 場が波数 l のモードに対して、どの程度の振幅を持つかを表す量になります。

波数 l のフーリエ係数は、 $\tilde{F}_l = |\tilde{F}_l| e^{i\theta_l}$ と表せるので、振幅と位相の 2 つの自由度を持ちます。宇宙原理の統計的等方性から、位相は興味のない量でないことが予想できるでしょう。⁴ そこで、振幅 $|\tilde{F}_l|$ の典型的な大きさを特徴づける統計量として、パワースペクトルを定義します。

$$\hat{P}_F(l) \equiv \frac{1}{N_{\text{mode}}(l)\Omega_s} \sum_{|l'| \in l} |\tilde{F}_{l'}|^2 \quad (2)$$

$\sum_{|l'| \in l}$ は、ビン幅内で $|l'| \in l$ を満たす全てのフーリエモードについての和とします。 $N_{\text{mode}}(l)$ は和に含まれる独立なフーリエモードの個数です。

$$N_{\text{mode}}(l) \equiv \sum_{|l'| \in l} \simeq \frac{2\pi l \Delta l}{(2\pi)^2 / \Omega_s} \quad (3)$$

Δl はビン幅 (観測者が決める量) です。 $2\pi l \Delta l$ は、フーリエ空間における半径 l 、幅 Δl の円環の面積 (2次元) であり、 $(2\pi)^2 / \Omega_s = l_f^2$ は基本波数で決まる面積要素です。この操作により、2次元データの膨大な情報量⁵を、 l の関数で与えられる1次元スカラー量のパワースペクトルに圧縮しています。宇宙原理に基づき、波数ベクトル l の方向依存性を無視し、大きさ $l = |l|$ を満たすフーリエ係数を全て同等と見なすことで、その振幅の平均がパワースペクトルです。フーリエ係数の振幅を定量化する統計量としては、最も単純なものと言えるでしょう。

地上の物理実験との違いは何でしょう? 地上実験の多くの場合、実験は何度も繰り返し行うことができます。多数の独立な実験結果の平均値と分散を取ること、実験結果の期待値と統計誤差を直接推定できるのです。しかし、宇宙論では観測領域は一つ、あるいは全天サーベイであっても宇宙は一つしかありませんので、これができません。このように、独立な事象 (リアライゼーション) によるアンサンブル平均を、観測データ内の同等の標本 (サンプル) 平均で代用する方

⁴ 異なる波数モード間でそろった位相が存在すると、結果の場 $F(\theta)$ には指向性が現れます。

⁵ 3次元データの場合も同じ。

法は、宇宙論におけるエルゴード仮定と呼ばれます。この仮定には、統計的不定性を伴うので、後に述べるように観測者がモデル化する必要があります。

3. 原始ガウシアンゆらぎと標本分散

前節では、宇宙論で良く用いられるパワースペクトルという統計量を定義しました。実は、パワースペクトルが観測データの「全て」の統計的情報量を記述する場合があります。そんな都合の良いことなんてあるの? と思われるかもしれませんが、あるのです。以下に見るように、宇宙は統計的にも単純で、美しいのです。

本記事では宇宙論のゆらぎ場に注目していますが、CMB温度ゆらぎ、宇宙の大規模構造など、宇宙に存在するゆらぎ (非一様性) の起源を説明するのがインフレーション宇宙シナリオです。ビックバン宇宙の始まりに (起こったと信じられている)、宇宙が指数関数的に膨張したというシナリオです。インフレーション時は、宇宙そのものも非常に小さかったので、インフレーション膨張を起こす場 (インフラトンと呼ばれる) の量子化を考える必要があります。量子力学の不確定性関係より、インフラトン場は必然的に量子ゆらぎを持ち、その量子ゆらぎがインフレーション膨張で引き伸ばされ、古典ゆらぎを生成したというのです。場の量子論では、異なる波数ベクトルのモードは異なる量子状態に対応します。通常のインフレーションモデルでは、インフラトン場への相互作用の影響は小さいと考えられ、異なる波数の量子ゆらぎはほぼ独立であったと考えられています。つまり、インフレーションによって生成された古典ゆらぎを $\tilde{\zeta}_{\mathbf{k}}$ (曲率ゆらぎと呼ばれる) とすれば、以下の条件が満たされることになります。

$$\langle \tilde{\zeta}_{\mathbf{k}} \tilde{\zeta}_{\mathbf{k}'} \rangle \equiv P_{\zeta}(k) (2\pi)^3 \delta_D^3(\mathbf{k} + \mathbf{k}') \quad (4)$$

$P_{\zeta}(k)$ は原始パワースペクトルです。 $\delta_D^3(\mathbf{k} + \mathbf{k}')$ は3次

元のデルタ関数で、異なる波数ベクトルのモードは独立であることを保証しています。また、インフレーション膨張が等方である限り、原始パワースペクトルの等方性、すなわち P_ζ が波数の大きさ $k = |\mathbf{k}|$ にも依存することも自然に帰結されます。ここで、記号 $\langle \rangle$ はアンサンブル平均を意味します。

このようにインフレーションは、ほぼ自由場の量子ゆらぎというランダム過程で、宇宙全体に等方的な古典ゆらぎを生成するのです。より正確には、古典ゆらぎ $\zeta_{\mathbf{k}}$ の位相は確率的な変数（ランダム）であり、その振幅は量子ゆらぎがどのくらい引き伸ばされたかという物理的な情報を含み、その典型的な大きさは原始パワースペクトル $P_\zeta(k)$ で与えられます。これは原始古典ゆらぎ場 $\zeta(\mathbf{x})$ がランダムガウシアン場であるということと等価です。ガウシアン場は非常に単純な統計的性質を持ちます。偶数次の多点相関関数⁶は2点相関関数（パワースペクトル）の積で与えられ、奇数次の多点相関関数⁷は場の対称性から、ゼロになります。つまり、**ガウシアン場の統計的性質はパワースペクトルで完全に記述されるのです。**

宇宙論の構造形成の問題とは、インフレーションなどで与えられたゆらぎの初期条件から出発し、輻射優勢期、⁸ 物質優勢期、⁹ そしてダークエネルギーが卓越する加速膨張期と変遷する膨張する宇宙のなかで、光子、バリオン、ダークマターなどの各成分のゆらぎの力学的進化を解くことです。ゆらぎの振幅が小さい限り、この多成分の力学進化は、アインシュタイン方程式とボルツマン方程式を組み合わせた力学系を「線形」摂動理論で調べることができます。線形解析では、異なる波数のフーリエ成分はそれぞれ「独立」に時間成長します。すなわち、ゆらぎが線形段階にある限りは、ゆらぎの統計的性質は原始ゆらぎのものを保持することになるのです。実際に、観測されるCMB温度ゆらぎはガウシアン場と矛盾しないことが示されており、インフレーションシナリオをサポートする一つの証拠となっています。¹⁰

このように、宇宙論データからパワースペクトルを

測定することは、インフレーション理論から動機づけられた自然なアプローチになります。しかし、前節で述べたように、有限観測領域からパワースペクトルを推定する際には、ゆらぎ場の有限モード数（標本）に起因する統計的不定性を考慮する必要があります。¹¹ この不定性は標本分散（sample variance）と呼ばれます。この統計的不定性は、推定パワースペクトルの共分散で与えられ、ガウシアン場については厳密に計算できます。

$$\begin{aligned} \text{Cov}[\hat{P}_F(l), \hat{P}_F(l')] &\equiv \langle \hat{P}_F(l) \hat{P}_F(l') \rangle - \langle \hat{P}_F(l) \rangle \langle \hat{P}_F(l') \rangle \\ &= \frac{2}{N_{\text{mode}}(l)} \delta_{ll'}^k P_F(l)^2 \end{aligned} \quad (5)$$

ここで $\delta_{ll'}^k$ はクロネッカーのデルタ関数で、ビン幅内で $l=l'$ のとき $\delta_{ll'}^k = 1$ 、それ以外は $\delta_{ll'}^k = 0$ と定義されます。このように、ガウシアン場の共分散行列は対角成分しか持ちません。言い換えれば、異なる波数ビンのパワースペクトルは独立であることを意味します。¹² 非ガウシアン場の場合には、パワースペクトルの積では表現できない4点相関関数の寄与を考慮する必要があります。一般に異なるビン間のパワースペクトルの相関が生じます。実際の測定では、検出器のノイズの影響等を考慮する必要がありますが、今回の主題ではありませんので、無視することにします。

上述の共分散行列は、ある面積 Ω_s の観測領域リアライゼーションから各波数 l のパワースペクトル $\hat{P}_F(l)$ を推定したときに、その測定値が期待値（真の値）まわりに分布するばらつき（ $\pm 1\sigma$ ）を与えます。つまり、統計誤差のことです。このことから、各波数ビンでのパワースペクトル測定の統計的有意度は、

⁶ 例えば、4点相関関数であれば、 $\langle \zeta(\mathbf{x}_1) \zeta(\mathbf{x}_2) \zeta(\mathbf{x}_3) \zeta(\mathbf{x}_4) \rangle$ 。

⁷ 例えば、3点相関関数であれば、 $\langle \zeta(\mathbf{x}_1) \zeta(\mathbf{x}_2) \zeta(\mathbf{x}_3) \rangle$ 。

⁸ 宇宙の全エネルギーに対して、光子、ニュートリノなど相対論的粒子が支配的な時期。

⁹ 宇宙が膨張するにつれ、相対論的粒子のエネルギー密度が減少し、ダークマターなど非相対論的粒子が宇宙の全エネルギーに対して支配的になる時期。

¹⁰ 例えば、天球上の各ピクセルで観測された温度ゆらぎの分布関数を調べると、その分布がガウシアンで非常に良くフィットできる。

¹¹ 宇宙の大きさは、我々の観測領域より大きいので、量子ゆらぎの波数刻みは、観測者のフーリエ分解能よりも常に高いと考えられます。つまり、観測するフーリエモードは多数の量子ゆらぎ起源の独立なゆらぎで構成されていると考えられます。

¹² 係数2は、ゆらぎ場の実数条件 $\hat{P}_F = \hat{P}_F^*$ という条件から、フーリエ係数の自由度が半分になることにより現れます。

$$\left[\frac{P_F(l)}{\sigma(P_F(l))} \right]^2 = \frac{N_{\text{mode}}(l)}{2} \quad (6)$$

と見積もれます。ここで、 $\sigma(P_F(l)) = \text{Cov}[(\hat{P}_r(l), \hat{P}_r(l))]^{1/2}$ です。このように、統計的有意度はパワースペクトルの値に依存せず、波数 l まわりのモード数 $N_{\text{mode}}(l)$ のみに依存することが分かります。 $N_{\text{mode}}(l) \propto \Omega_s l \Delta l$ であるので、観測領域 (Ω_s) が大きいほど、波数 l が大きいほど、またピン幅 Δl が大きいほどパワースペクトル測定の有意度は高くなります。

図1の右下図は、Planck衛星によるCMB温度ゆらぎのパワースペクトルの測定結果を示します。波数 l (厳密には球面調和関数の展開次数) が大きくなるにつれ、パワースペクトルの測定精度が向上しているのが分かります。各 l ビンのパワースペクトルの測定値まわりの誤差棒は、上述した標本分散と検出器のノイズの寄与を考慮したものです。

パワースペクトル推定の統計誤差が与えられれば、理論モデルと比較することができます。

$$\hat{P}_F(l) \longleftrightarrow P_F^{\text{model}}[l; P_c(k), \Omega_{\text{m}0} h^2, \Omega_{\text{b}0} h^2, \Omega_{\text{de}}, \dots] \quad (7)$$

右辺は、理論モデルによる F 場のパワースペクトルが、原始パワースペクトル、その他宇宙膨張を記述するための宇宙論パラメータの関数として与えられることを意味しています。これはゆらぎが線形成長する限り、正しい仮定です。まず、統計誤差内で、理論モデルが観測データを再現することができるかどうかを調べます (図1の右下図参照)。次に、統計誤差が許す範囲をパラメータ決定に正しく伝播させることで、各パラメータの決定の信頼区間 (C.L.) を求めることができます。Planckチームは、このような作業過程を用いて、ダークマター、バリオン、宇宙年齢などのパラメータを高精度で決定したのです。図1の左下図に示されているのは、そのような例の一つです。特筆すべきは、Planckの場合には、すでに全天データが存在し、また $l \sim 2000$ までのデータについては、検出器に

よる誤差が効かない、標本分散による統計精度まで実験結果が達成していることです。つまり、人類は(温度ゆらぎについては) CMBデータから引き出すことが可能な全ての統計情報を宇宙論に用いることに成功したのです。

以上ここまでの話をまとめると、(1) 宇宙原理の統計的一様性、等方性を仮定し、興味あるゆらぎ場の統計的性質を測定するために、一次関数であるパワースペクトルを測定すること、(2) エルゴード仮説、つまり観測した領域が宇宙の典型的な標本であると仮定し、パワースペクトルの測定に伴う標本分散をモデル化すること、(3) インフレーションシナリオが予想するように、(線形段階の) ゆらぎ場がガウシアンである場合、ゆらぎ場の統計的性質はパワースペクトルで全て決定されること、を述べてきました。これら (1), (2), (3) のいずれかの仮定が破れれば、パワースペクトル以外の統計量を用いる必要があることになります。

4. 宇宙の構造形成: 重力の非線形性

前節までは、CMBゆらぎ場に代表される、線形かつガウシアンゆらぎ場を考えてきましたが、この節では加速膨張する宇宙の探査を目的とする銀河サーベイから得られる宇宙論データを考えてみましょう。CMB以降の物質優勢期のゆらぎの力学進化は、主にダークマターの重力により引き起こされます。重力の不安定性により、ダークマターの空間分布の非一様性が増幅され、現在観測される、銀河、銀河団、さらに銀河の分布で見られる宇宙の大規模構造が形成されたというシナリオです (以後CDM構造形成モデル)。この CDM構造形成モデルでは、より小スケールの構造から形成され、徐々に大きなスケールの構造が形成されたという、ボトムアップ的な階層構造形成シナリオを予言します。

冷たいダークマターの仮定は、その熱的速度が小さい (冷たい)、重力のみで相互作用する、という性質です。ホライズン内の領域で、また空間的に粗視化し

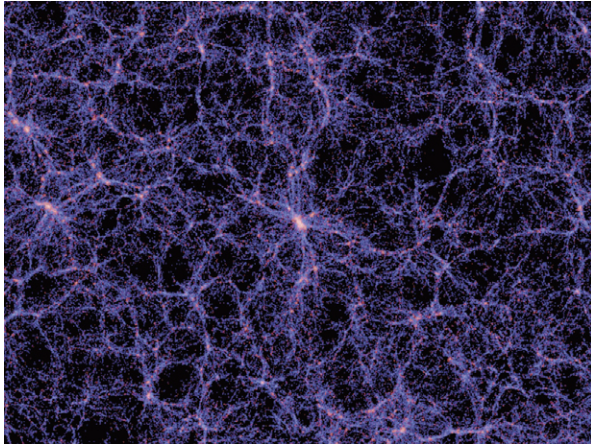


図2 冷たいダークマター構造形成モデルに基づく宇宙論 N 体シミュレーションの例。ガウシアン場の初期条件から出発したとしても、非線形重力進化の結果として、ダークマターの分布は複雑な非ガウシアン性を有する。ダークマターが特に集中している領域はダークマターハローと呼ばれる。フィラメントの交差点には、太陽質量の 10^{15} 倍もの銀河団スケールの巨大ハローが存在することがある（図の中心にある巨大ハローはそのような例）。宇宙全体の総質量に対して銀河スケール以上のハローに含まれるダークマター質量は数10%にも及ぶ。一方、70%ほどの体積比は、ダークマター密度が平均より少ない、質量密度ゆらぎが負になっているボイド領域が占める。このような非対称性より、ダークマター分布は、ゼロでない多点相関関数を持つことになる。

た観点では、ダークマターの質量密度場、速度場は、膨張宇宙における渦なし、無圧力の流体の方程式系に従うことが示されます。¹³

$$\begin{aligned}\frac{\partial \delta_m}{\partial t} + \frac{1}{a} \nabla \cdot [(1 + \delta_m) \mathbf{v}_m] &= 0 \\ \frac{\partial \mathbf{v}_m}{\partial t} + \frac{\dot{a}}{a} \mathbf{v}_m + \frac{1}{a} (\mathbf{v}_m \cdot \nabla) \mathbf{v}_m &= -\frac{1}{a} \nabla \phi \\ \nabla^2 \phi &= 4\pi G \bar{\rho}_m a^2 \delta_m\end{aligned}\quad (8)$$

$a(t)$ は宇宙のスケール因子であり、宇宙膨張とともに増加する関数、 $\delta_m(\mathbf{x}) \equiv [\rho_m(\mathbf{x}) - \bar{\rho}_m]/\bar{\rho}_m$ は質量密度ゆらぎ場であり、 $\mathbf{v}_m(\mathbf{x})$ は固有速度ベクトル場、 $\phi(\mathbf{x})$ は重力ポテンシャルです。一様等方宇宙では、至るところで $\delta_m = |\mathbf{v}_m| = 0$ になる宇宙ですので、 δ_m と \mathbf{v}_m はゆらぎ場です。CMB の測定で制限されているゆらぎの初期条件から出発し、この方程式系を解くことにより、宇宙構造の力学進化を調べることができます。この方程式系から明らかのように、ゆらぎ場の振幅が小さいとき、つまり $|\delta_m| = |\mathbf{v}_m| \ll 1$ (光速 $c=1$ の単位系) では、方程式は線形化でき、ゆらぎ場は線形進化します。しかし、ゆらぎが時間とともに成長し、非線形項 ($\delta_m \mathbf{v}_m$ 、 $(\mathbf{v}_m \cdot \nabla) \mathbf{v}_m$) が無視できなくなると、ゆらぎ場は非線形進化することになります。つまり、異なる波数のフリーモードが混合し (モードカップリング)、複雑な

¹³ 厳密には無衝突ボルツマン方程式系を解く必要があり、N体シミュレーションはそれを近似的に解く方法になっています。

進化を始めるのです。原始ゆらぎがガウシアン場があったとしても、重力の非線形性がダークマターの空間分布にガウシアン性を誘発するのです。その非ガウシアン性の程度は、小スケールほど、また現在に近い低赤方偏移ほど大きいこととなります。

このように、現在の宇宙のダークマターの分布の統計的性質は、パワースペクトルの情報だけでは記述できません。実際に図2に示されるような、CDM構造形成モデルのN体シミュレーションの研究は、ダークマターの分布が一般に3点以上の多点相関関数の値を持つようになることを示しています。例えば、非ガウシアン性の情報を持つ最低次の3点相関関数を考えてみましょう。非線形構造の成長の結果として、低密度領域については、最小でもダークマターが空っぽの領域 ($\rho_m(\mathbf{x}) = 0$) になりますが、そこでは $\delta_m(\mathbf{x}) = -1$ です。一方、ダークマターが密集する領域では、質量密度 $\rho_m(\mathbf{x})$ は幾らでも増幅する可能性があり、実際にN体シミュレーションではダークマターが密集するダークハローの中心で密度が発散する領域 ($\delta_m \rightarrow \infty$) が現れることを予言しています。このように、密度分布の非対称性により、一般に3点相関関数が値を持つようになるのです。

以上を踏まえ、宇宙論場の統計的情報量という観点から疑問が生じます。前節で述べたように、線形段階にある宇宙初期ゆらぎはガウシアン場であり、その統

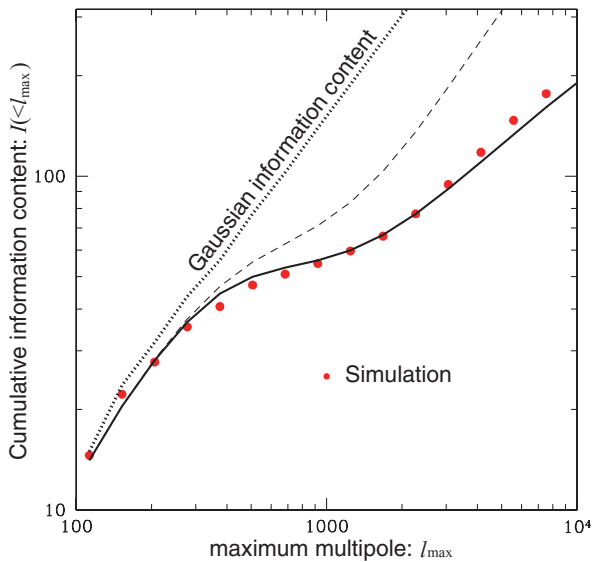


図3 宇宙の大規模構造による重力レンズ場のパワースペクトルに含まれる情報量 $I(<l_{\max})$ 。ここで考える重力レンズ場とは、観測者と赤方偏移 $z_s = 1$ にある光源銀河間にある質量密度ゆらぎ場を視線方向に投影した2次元場。観測面積として1400平方度を仮定し、パワースペクトル推定の統計誤差として、標本分散のみを考慮した（銀河固有の楕円率によるノイズは無視した）。示すのは、各波数ビンのパワースペクトル推定の統計的有意性（ガウシアン場については式6を参照）を、最小波数 $l_{\min} = 72$ から、 x 軸に与えられる最大波数 l_{\max} まで積分した量。この場合、 $I(<l_{\max})$ はピン幅 Δl に依存しない。点線カーブは、重力レンズ場がガウシアンの場合の結果、つまり宇宙の大規模構造の初期線形ゆらぎ場が持つ最大情報量。この場合、情報量はモード数の和にだけ依存し（式6）、 $I_{\max} \propto l_{\max} \Omega^2$ と表せ、 l_{\max} までのフーリエ空間での面積の平方根に比例する（3次元場の場合には、 $I(<k_{\max}) \propto k_{\max}^3 V_s^{1/2}$ となる。 V_s は観測量の3次元体積）。丸点は、図2に示されるようなCDM宇宙構造形成モデルのN体シミュレーションの結果を用い、重力レンズ場を再現した疑似カタログを使った結果。この場合、情報量を求めるために、重力レンズ場の非ガウス性のために生じる異なる波数のパワースペクトル間の相関を正しく考慮している。実線は解析的なモデルで、模擬カタログの結果を良く再現している。破線は、観測領域を超えるゆらぎ場による標本分散への影響を無視した場合（詳しくは本文参照）。

計的信息はパワースペクトルが全てです。一方、現在の宇宙では、ダークマターのゆらぎ場は非ガウシアン性を有し、パワースペクトル（2点関数）以外に多点相関関数が値を持つようになります。熱力学第二法則的な考えから初期条件以上の情報を引き出すことはできませんので、ダークマターの統計量の時間進化を考えると、非線形構造形成の結果として、ガウシアン場のパワースペクトルの一部の情報が多点相関関数の情報へ流出したとも考えられます。つまり、疑問は

- 現在のダークマターの空間分布から得られる統計量を組み合わせることで、宇宙初期に持っていたガウシアン情報量を復元できるか？

ということです。実は、これは宇宙論の業界で**未解決問題**になっています。力学系が時間可逆な系であれば、完璧に初期条件に戻せるので、復元は可能でしょう。CDM構造形成モデルでは、大スケールの密度ゆらぎは、まだ線形段階にあり、ガウシアン性を保っているため、復元は可能ということになります。一方、小スケールの強非線形領域にある密度ゆらぎは、初期条件の情報を既に失っているかもしれません。例えば、ダー

クマターハロー内で重力的に束縛されているダークマターは、何度もハロー中心回りを振動あるいは散乱してきた軌道を持っていた可能性があり、現在の粒子分布の情報だけでは、可逆的に初期条件に戻すことは不可能なように思います。これら中間スケールの弱非線形領域では、重力進化がまだ十分に進んでおらず、初期条件のガウシアン情報の「大部分」を復元することが可能かもしれません。「大部分」と書きましたが、どの程度復元できるかについては、まだ良く分かっていません。実は、銀河サーベイから測定できる、銀河のクラスタリング、重力レンズの宇宙論統計量の興味あるスケールはこの弱非線形領域にあります。

この数年、私たちは構造形成のN体シミュレーションあるいは解析的モデルを用いて、この問題を調べてきています。図3は、その研究から得られた結果の一つです。宇宙の構造形成のN体シミュレーションを使って再現した宇宙の中を光を飛ばし、重力レンズ効果の観測量を復元した疑似カタログを用い、その重力レンズ場のパワースペクトルがどの程度の情報量を持っているかを調べた結果です。波数 l が小さいところでは、この領域は線形段階のゆらぎの影響が支配的ですが、ガウシアン場から期待される情報量をほぼ復

元しています。一方、波数 l が数100以上のところでは、パワースペクトルが持つ情報が大きく減少していることが分かります。HSCのような今後のサーベイが宇宙論に用いる、1～1000のスケールでは、半分以上の減少があります。この結果は、パワースペクトル解析だけではガウシアン情報の一部しか復元できないことを示しています。実際に、別の研究では、3点相関関数が情報を加えることを見つけていますが、それでもまだガウシアン情報を復元することには成功していません。4点以上の相関関数も同様に重要かもしれません。

これらの研究を通して、予想しなかった面白い結果も発見しています。実は、ダークマターのパワースペクトルにおける情報の消失は、主にサーベイ領域より大きなゆらぎによって引き起こされているのが分かったのです。観測領域を超えるゆらぎは観測できないため、観測領域が、一様等方宇宙より正の密度ゆらぎ領域にいるのか、あるいは負の密度ゆらぎ領域にいるのか分からないのです。つまり、観測領域内で平均した密度ゆらぎは一般にゼロではありません。重力の遠距離力、非線形性の性質により、全ての波長はモードカップリングしますので、もし観測領域が正の密度ゆらぎの領域にいれば、ダークマターの質量密度パラメータが宇宙全体の平均より若干大きな宇宙、つまり正の曲率の宇宙にいることと同等になり、観測領域内の全てのスケールのゆらぎの成長が加速されることとなります。もし観測領域が負の密度ゆらぎの領域であれば、小スケールのゆらぎの成長は抑制されるのです。実は、このモードカップリングによる影響が標本分散に最も寄与することが分かり、我々はこの効果を定式化することに成功しました。図3の実線は、この効果を考慮した予言で、N体シミュレーションの結果を良く再現していることが分かります。逆に、この効果をパラメータに推定に正しく考慮することで、大スケールのゆらぎを制限できる可能性があります。これは非常に面白い可能性で、今後もさらに研究を進めていく予定です。

5. 今後の課題

IPMUが進めるSuMIRe計画に代表されるように、今後の宇宙論銀河サーベイは益々大型化し、またダークエネルギー問題に代表される、より難しい、より根源的な問題に挑むこととなります。この記事では、宇宙論観測データから統計量を測定し、宇宙論パラメータを推定するときの作業過程、その前提になっている原理、仮定を解説してきました。CMBの場合は、ゆらぎ場がほぼガウシアンであるために、最大限の成功を収めています。一方、銀河サーベイの場合には、重力の非線形進化の帰結として、どんな統計量が最適であるのか未だ分かっていません。また、有限領域に起因する標本分散をモデル化する困難さもあります。逆に言えば、未解決問題が沢山あり、嫌というほど研究の余地があるのです。銀河サーベイによる宇宙論をCMB宇宙論のレベルまで成熟させるのが目標と言えます。今回は、宇宙論の統計的側面に焦点を当てましたが、今後もこの流れは益々強くなると思われます。新しいアイデアがある方、是非私の方までご連絡ください...特に統計が専門の方、ご連絡お待ちしております!

参考文献

- M. Takada and S. Bridle, *New J. Phys.* **9**, 446 (2007)
- M. Takada and B. Jain, *MNRAS* **395**, 2065 (2009)
- M. Sato et al., *Astrophys. J.* **701**, 945 (2009)
- I. Kayo, M. Takada, and B. Jain, *MNRAS* **429**, 344 (2013)
- M. Takada and W. Hu, *Phys. Rev. D* **87**, 123504 (2013)
- Y. Li, W. Hu, and M. Takada, *Phys. Rev. D* **89**, 083519 (2014)
- M. Takada and D. N. Spergel, *MNRAS* **441**, 2456 (2014)