J. Plasma Fusion Res. Vol.85, No.10 (2009) 665‐673

講座 流体乱流研究から診たプラズマ乱流データの解析

3．確率密度関数とその応用

辻　義之，田中宏彦，大野哲靖
名古屋大学大学院工学研究科

（原稿受付：2009年 8 月16日）

本章では，変動の大きさの分布を特徴づける確率密度関数の定義と計算方法を述べたのち，変動特性の異なるプラズマ揺動データに適用して視覚的に理解を促す．流体乱流中で見られる代表的な分布を示し，その定量的な評価法や直交展開，分布を比較する手法などを紹介する．最後に，複数の確率変数から計算される結合確率密度関数について述べる．

Keywords:
statistical analysis, fluctuation, probability density function, distribution, moment, Gram-Charlier expansion, Kullback-Leibler divergence, joint probability density function

3.1　はじめに

大学院の講義に，「熱」に関する多様な側面を体系的に講義する科目がある．分子・原子論的な観点から始まり，巨視的な熱力学を学び，熱の移動は微分方程式（熱伝導方程式）によって決定論的に記述されることを理解する．しかし，実際に実験などで計測される熱は，時間・空間的に揺らいでいるものであり，その揺らぎを客観的に特徴づけるためにはどのようにすればよいのですか？と学生に問うと困惑してしまうようである．この揺らぎに関する時間的な性質を特徴づけるのがスペクトルであり，揺らぎの大きさの分布を解析する手法が，確率密度関数である．本章では確率密度関数について解説する．

確率論を学ぼうとすると，まず確率空間，測度，確率変数といった抽象的な概念の理解からはじまる[1]．それは重要であることに変わりないが，初めて学ぼうとする者にとっては大きなハードルになるであろう．本章では，細かな定義に深入りして議論を煩雑にすることは避け，具体的な解析例とその解釈を示すことに重点をおいた．解析に用いたデータは第二章と同様にプラズマ中で計測されたイオン飽和電流を対象にしている．乱流は私たちの日常に広く見られる現象であり，古典力学で最も難解な問題と考えられている．確率密度関数は，過去の長い歴史において流体乱流（中性流体）の実験や観測のデータの統計処理法として広く利用されてきた．確率密度関数を計算し，そのモーメントを求めること，また確率密度関数型の特徴づけについて解説したい．限られた内容ではあるが，これらが，今後は広くプラズマ乱流のデータ解析に利用され，現象の新たな理解に役立つことを期待するものである．

3.2　確率密度関数

まず，確率変数について説明したい．例えば，サイコロを 1 回投げると，その数は 1 から 6 の値のどれかを必ず出す．しかし，どの値がでるのかを事前に知ることはできない．このように，どの値（もしくは，どういう範囲の値）が生起するかが事前にわかってはいるが，実際にどのような値が実現されるかは，確率に支配されているような関数を確率変数と呼ぶ．計測などで得られる時系列データは確率変数とみなされ，その分布を知るために確率密度関数は必須となる．

3.2.1　定義と具体的な計算方法

確率変数 x がある任意の実数 ξ（これを状態変数と呼ぶ）より小さくなる確率 $P_x(\xi)$ を確率分布関数（probability ditribution function）と呼ぶ．すなわち，記号 prob[ ] をカッコ内の条件を満たす確率を表すこととすると，

$$P_x(\xi)=\mathrm{prob}[x\leq\xi], \tag{1}$$

が確率分布関数である．$P_x(\xi)$ は次の 3 つの性質を満たす：(i) $P_x(\xi)\geq 0$，(ii) $P_x(+\infty)=1$，$P_x(-\infty)=0$，(iii) $\eta>\xi$ のとき $P_x(\eta)>P_x(\xi)$．確率分布関数 $P_x[\xi]$ がいたるところ連続で微分可能な場合には，その微分が計算され

$$p_x(\xi)=\mathrm{d}P_x(\xi)/\mathrm{d}\xi, \tag{2}$$

となる．ここで，$p_x(\xi)$ を ξ の確率密度関数（Probability Density Function；以降は PDF と略記する）と呼ぶ．また，確率の定義より $p_x(\xi)$ は非負の実関数であり，次の性質をもつ：(i) $p_x(\xi)\geq 0$，(ii) $\int_{-\infty}^{+\infty}p_x(\xi)\mathrm{d}\xi=P_x(+\infty)=1$，(iii) $\int_a^b p_x(\xi)\mathrm{d}\xi=P_x(b)-P_x(a)$．

実験や観測で収集される離散データを取り扱う場合には，分布関数を求めてその微分から確率密度関数を算出することはあまり行われない．それは，数値微分をする際の誤差が無視できないからである．**図 1** にはプラズマ乱流中で計測されたイオン飽和電流の時系列信号を示す．ここで

3. Probability Density Function and Its Application to Statistical Analysis
TSUJI Yoshiyuki, TANAKA Hirohiko and OHNO Noriyasu
corresponding author's e-mail: ohno@ees.nagoya-u.ac.jp

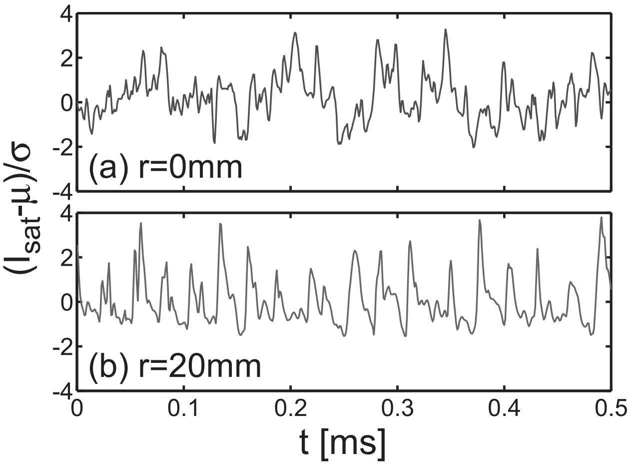

図1 NAGDIS-II の接触プラズマ中における静電プローブ（ラングミュアプローブ）計測によるイオン飽和電流波形．r はプラズマ中心からの距離．

信号は平均値と標準偏差値で無次元化してある．このデータを用いて，PDF を実際に計算してみたい．具体的なイメージを持てるように，確率変数 $x(t)$ が時間 t の関数として観測されることとする．この場合，$x(t)$ が ξ と $\xi+\varDelta\xi$ の間にある時間の区間和 $\sum_i \varDelta\tau_i$ の全区間 T に対する割合，すなわち

$$\mathrm{prob}[\xi < x(t) < \xi+\varDelta\xi] = \lim_{T\to+\infty}\sum_i \varDelta\tau_i/T, \tag{3}$$

を求める．実際の計測では，時間間隔 $\varDelta t$ でサンプリングされた離散データ点を扱うため，式(3)は離散点が，$(\xi,\xi+\varDelta\xi)$ の範囲に含まれる割合を数えることに等しくなる（**図2**参照）．測定時間 T 内の全データ点数を N $(=T/\varDelta t+1)$，含まれるデータ点の個数を n_j とすると，PDF は $\varDelta\xi\to 0$ の極限をとることにより，

$$p_x(\xi) = \lim_{\varDelta\xi\to 0}\lim_{T\to+\infty}\sum_i \varDelta\tau_i/T = \lim_{\varDelta\xi\to 0}\lim_{N\to+\infty}\sum_j n_j/N, \tag{4}$$

と計算される．しかし，実際の実験では計測時間 T は有限であるし，$\varDelta\xi$（この大きさを bin size と呼ぶ）をゼロの極限に持っていくこともできない．計測時間 T をどれだけ長くとればよいのか？という疑問は，必ず誰もが抱くはずであるが，これに関してはPDFのモーメントに関連させて後に述べる（3.2.4節）．bin size の大きさは，経験的に $x(t)$ の標準偏差 σ の10%～20%が妥当ではないかと思う．**図3**は例に掲げた信号のPDFである．横軸は平均値 μ と標準偏差 σ で無次元化してある（平均値と標準偏差の計算方法は後述する）．これによって，変動の大きさが標準偏差の何倍であるかが明確に理解できる．bin size を標準偏差の20%とすることは，**図3**の横軸が0.2きざみで確率の大きさが計算されることに対応する．このように，標準偏差を基準に bin size を設定することは，離散データの統計処理においても有利である．十分に計測時間がある場合には，bin size を変化させても PDF 形状が大きく変化しないことが，計算結果の妥当性の指標となる．

3.2.2 代表的な PDF の例

流体乱流の解析でしばしば出くわす確率密度関数の代表的な例を以下にまとめた．**図1**に示した具体的例との対応

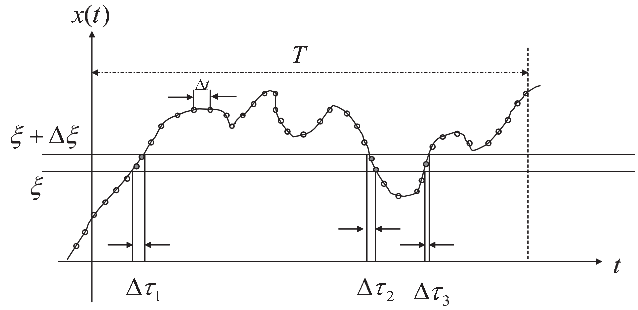

図2 確率密度関数の計算方法の概念図．

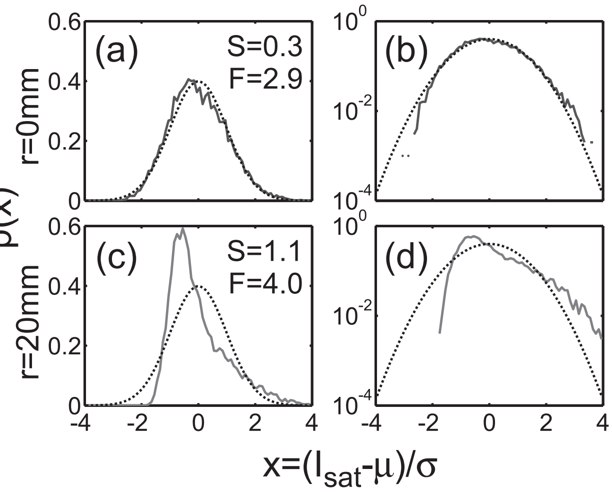

図3 図1の信号波形から計算された確率密度関数．左カラムの縦軸は linear，右のカラムの縦軸は log．図中にゆがみ度 S と偏平度 F の値が示してある．

からも理解をすすめてもらいたい．なお，標記を簡略化するため，以降は PDF を $p(x)$ とあらわすこととする．

［正規分布（ガウス分布）］

不規則変動の分布として広くみられ，流体乱流では一様等方場で観測される速度変動はこの分布に近い．**図1**に示したイオン飽和電流の変動が正規分布に近いと考えられる．また，LHD での解析例は，文献[2]を参照されたい．

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]. \tag{5}$$

ここで，μ は平均値，σ は標準偏差をあらわす．

［指数分布］

流体乱流では，速度変動の微分や粒子加速度のPDFのすそ野がこの分布に従う[3]．**図1**(b)の確率のすそ野が指数分布に近いと考えられる．

$$p(x) = a\exp(-bx). \tag{6}$$

ここで，$a>0$ として，$b>0$，$x\geq 0$ もしくは $b<0$，$x\leq 0$ とする．流体乱流では，その間欠的な性質から現れる，確率分布のすそ野を $p(x)\propto\exp(-|x|^s)$ としてよりよく近似することもおこなわれる．この分布を引き伸ばされた指数分布（stretched exponential）と呼ぶことが多い[3]．

［ベキ乗分布］

乱流速度変動のレベル交差問題の待ち時間の分布が従

う．ベキ乗分布では，分布の成り立つ範囲$[x_1 \leq x \leq x_2]$が必ず存在する．この最小，最大スケールは，物理スケールと対応している場合が多いので，ベキ指数とあわせて明示するのがよい．フラクタル分布の関連として解釈されることもある[4]．

$$p(x) = ax^{-b}. \tag{7}$$

ただし，$a>0$，$b>0$とする．

［対数正規分布］

確率変数の対数をとった変数が正規分布に従う．流体乱流のエネルギー散逸の大きさは，この分布でよく近似される[5,6]．プラズマ乱流中での解析例は，例えば文献[7]にある．

$$p(y) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right], \quad y = \log(x), \tag{8}$$

ただし，μとσはyの平均値，標準偏差をあらわす．

確率分布関数型を特徴付ける際に注意すべき点は，**図3**左列のようにPDFの大きさをそのまま表示した場合と，**図3**右列のように対数表示した場合である．正規分布の場合であっても，平均値まわりで生起する現象を対象とする場合には，そのままの表示でよいが，まれに起こる現象が問題になる場合には，対数表示をして確率のすそ野の形状を明瞭にして議論するほうがよいであろう．その他にもいくつかの重要な確率分布型があるので，詳しくは参考文献[8]を参照いただきたい．また，プラズマ乱流中での解析例は文献[9,10]などにある．

3.2.3 特性関数

確率過程に関する多くの情報を圧縮して含む有効な関数として特性関数がある．特性関数は確率密度関数のフーリエ変換として定義される[11]．

$$\phi(u) = \int_{-\infty}^{+\infty} \exp(iux)p(x)\mathrm{d}x, \tag{9}$$

$$p(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \phi(u)\exp(-iux)\mathrm{d}u. \tag{10}$$

特性関数は，中心モーメントや確率密度変数の和を計算するのに有効に利用される．実験データを取得する際には，時間連続（アナログ）データをデータロガーなどでサンプリングして離散（デジタル）データとして保存される．したがって，PDFを計算する過程では離散化による誤差が含まれる．それをさらにフーリエ変換することは，またいくつかの制約を課すことになる（詳しくは，第2章スペクトルの解説を参照）ので，乱流の実験データを解析する限りは，特性関数を求めることはあまりないようである．

3.2.4 中心モーメント

確率変数の分布特性として，値の拡がりや対称性を調べるためにはPDFを積分して情報を圧縮した数値を指標にすると便利である．平均値μや標準偏差σは，1次，2次モーメントとして以下に与えられる．

$$\mu = \int_{-\infty}^{+\infty} xp(x)\mathrm{d}x = \sum_{i=1}^{N} X_i/N. \tag{11}$$

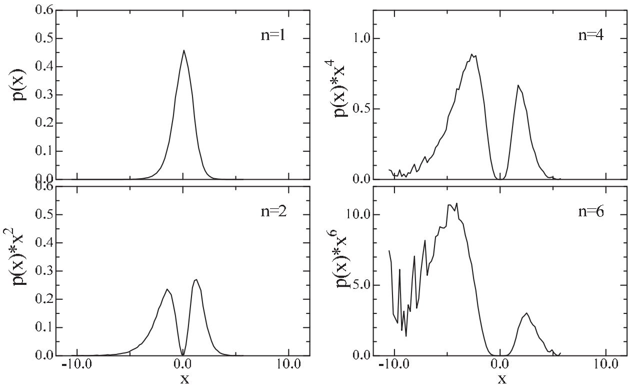

図4　高次モーメントを計算する際のPDF型の変化．

$$\sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 p(x)\mathrm{d}x = \sum_{i=1}^{N} (X_i-\mu)^2/N. \tag{12}$$

ここで，X_iは観測される離散データをあらわす．μは分布の重心であり，期待値とも呼ばれる．標準偏差の2乗（σ^2）は分散と呼ばれ，重心まわりの分布の広がりを表す指標となる．流体乱流の解析では，平均値まわりの変動の様子を主に調べることが多いので，平均値まわりのモーメントを調べることが多く，中心モーメントと呼ばれる．

分布の偏りやすそ野の広がりは，以下に示すn次の中心モーメントを計算することで特徴づけられる．

$$m_n = \int_{-\infty}^{+\infty} (x-\mu)^n p(x)\mathrm{d}x = \sum_{i=1}^{N} (X_i-\mu)^n/N. \tag{13}$$

ゆがみ度（skewness）Sは，xの3次の中心モーメントを2次の中心モーメントで正規化したものである．

$$S = m_3/\sigma^3. \tag{14}$$

確率変数の3乗を計算するために，その平均値からのズレが符号により反映され，分布の対称性を見積もる指標となる．Sが負であればマイナス側，正であればプラス側にPDFのすそ野が広がっていることが予想される．一方，偏平度（flatness）Fは，xの4次の中心モーメントを2次の中心モーメントで正規化したもので，次式で定義される．

$$F = m_4/\sigma^4. \tag{15}$$

xの4乗を計算するために符号の情報は反映されないが，確率変数の大きな値（平均値から離れた値）がFには反映されることとなる．したがって，分布の広がり（PDFのすそ野）のなだらかさを定量化する指標と理解できる．この数値が大きいほど，分布のすそ野がなだらかで広がっていることを表す．**図3**に示したPDFには，SとFの値を示してある．PDFの型とこれらの数値との対応を比較してもらいたい．なお，正規分布の偏平度は3であることから，3からのズレは分布のすそ野が正規分布に比べてどの程度の相違があるかを定量的に表す指標になる．プラズマデータの解析例としては，文献[12]などがある．

確率変数のn次モーメントは，式(13)の積分核$(x-\mu)^n p(x)$がx軸とで囲まれる部分の面積に対応する．

よほど特殊な分布でない限り，$(x-\mu)^n p(x)$ は x が十分に大きなところでゼロに漸近する．しかし，実際にデータを解析してみると，必ずしもそのようにはならないことがある．**図4**にその一例を示す．指数 n を 1，2，4，6 と変化させた場合の積分核 $(x-\mu)^n p(x)$ を示した．$n=6$ の場合にはグラフはゼロに漸近せず，確率分布のすそ野が正確に計測できていないことを示唆している．この結果を用いて計算される m_6 は，過大もしくは過小評価となるおそれがある．したがって高次のモーメントを正確に計算するためには，各モーメントの次数 n に対して**図4**の分布をチェックして，確率のすそ野まで正確に計測できているかを（つまり，積分核 $(x-\mu)^n p(x)$ が x の大きなところでゼロに漸近することを）確かめる必要がある．そのためには，計測時間 T を十分にとって確率の小さい事象に対しても，十分なサンプル数が得られるようにする．これが先に述べた計測時間 T に対する一つの指標である．ただし，確率のすそ野を正確に計測することは，計測時間とは独立に測定器の精度やノイズとの関連で一概には改善できない問題も含んでいる．

PDF は統計量として基礎的であり，かつ重要なことは改めて述べる必要はない．離散データからPDFを計算する際に注意することは，bin sizeと計測時間 T であることを説明した．しかし，確率密度関数をいくら正確に調べても，確率変数の統計的性質のうちどうしても見出すことができない性質がある．それは，時間に関する情報である．確率密度関数には，その算出過程からも分かるように（**図2**），時間軸は縦軸上に射影されるため，同一の確率分布型を持ちながら，異なる時間波形を無数に作ることが可能となる．確率変数の時間情報を知るためには，第2章の周波数スペクトルもしくは自己相関関数を調べる必要がある．また，計測時間 T に関しては，スペクトルが正確に計算できるための基準からも触れているので参照されたい．

3.3 確率密度関数型の表現

PDF を計算した場合に，その分布型の特徴を説明することは容易なことではない．幾つかの PDF の分布を比較して，物理現象との対応を定性的に推測することも可能であろう．3次や4次のモーメントは確かに有用な情報ではあるが，より定量的な記述ができるに越したことはない．本章ではより高度な PDF 型の特徴づけについてまとめる．

3.3.1 PDF 型の特徴づけ

流体乱流の解析で，PDF 型を特徴づける方法は過去にいくつかが試みられている．その中で Barndorff-Nielsen は下記のような一般型を提案している[13]．

$$p(x)=a\exp\left\{-\frac{b\left[\sqrt{c^2+(x-d)^2}-(x-d)\right]^e}{2}-\frac{f\left[\sqrt{c^2+(x-d)^2}+(x-d)\right]^g}{2}\right\}, \quad (16)$$

ここで，a は規格化定数，b，c，d，e，f，g は任意定数である．式(16)を用いると正規分布，双曲分布，指数分布やその混合型が表現できる．測定から得られた PDF 型に式(16)をフィットして，パラメータ $a \sim g$ を求め，実験条件を変えた場合にこれらの定数がどのように変化するのかを調べれば，PDF 型の変化をより定量的に議論できる．乱流解析では PDF 型のすそ野が指数型になる場合があり，その指数の（例えばレイノルズ数に対する）変化を調べることがある．式(16)で $x \to +\infty$ の場合を考えると，$p(x) \simeq a\exp\{-f(x-d)^g/2^{1-g}\}$ となるので，指数 g の変化に重点を置いて調べれば良いことになる．乱流での解析例は，文献[14]を参照されたい．

3.3.2 PDF の直交展開

確率密度関数型が正規分布に近い場合には，正規分布からのズレを補正項として表すことができる．この補正項はエルミート（Hermite）多項式により記述され，Gram-Charlier 級数展開と呼ばれる．確率変数 ξ をその平均と標準偏差により無次元化した変数 $x=(\xi-\mu)/\sigma$ を考える．正規分布を $\phi(x)$ とすると，確率密度関数は次のように展開できる．

$$p(x)=c_0\phi(x)+\frac{c_1}{1}\phi^{(1)}(x)+\frac{c_2}{2}\phi^{(2)}(x)+\cdots+\frac{c_n}{n!}\phi^{(n)}(x), \quad (17)$$

$$\phi^{(n)}(x)=(-1)^n H_n(x)\phi(x), \quad (18)$$

ここで，$H_n(x)$ はエルミート多項式であり，

$$\begin{aligned}&H_0(x)=1, \quad H_1(x)=x, \quad H_2(x)=x^2-1\\&H_3(x)=x^3-3x, \quad H_4(x)=x^4-6x^2+3\\&H_5(x)=x^5-10x^3+15x, \ \cdots\end{aligned} \quad (19)$$

次のような直交性を備えている．

$$\int_{-\infty}^{+\infty} H_m(x)H_n(x)\phi(x)\,\mathrm{d}x=\delta_{mn}(x)=\begin{cases}1 & \text{for } (m=n)\\ 0 & \text{for } (m\neq n)\end{cases} \quad (20)$$

したがって，式(17)は確率密度関数の直交展開と解釈することができる．展開係数 c_n は直交性を利用して，

$$c_n=(-1)^n\int_{-\infty}^{+\infty} H_n(x)p(x)\,\mathrm{d}x, \quad (21)$$

から求められ，$c_0=1$，$c_1=c_2=0$，$c_3=-m_3/\sigma^3=-S$，$c_4=m_4/\sigma^4-3=F-3$ となり，高次の係数も中心モーメント（m_n）の関数として一意に定まる．$p(x)$ が正規分布の場合には，3次以降の係数はすべてゼロとなる．それゆえ，確率密度関数型の正規分布からの偏差は各係数の分布から定量的に評価される．**図1**の信号波形を用いて直交展開した結果を**図5**に示す．また，展開係数の分布は**図6**に示した．流体乱流での解析例は，文献[15]を参照されたい．

3.3.3 条件付き統計量による PDF 型の表現

計測データから確率密度関数を計算してその変化を調べることは，現象を理解する上で有意義なことであろう．しかし，より詳しい議論をするためには，PDF 型がどのような物理量に関連して決まるのかがわかるとよい．以下では，その1つの方法を紹介したい[16]．

定常な確率変数 $x(t)$ に対して，その微分（$\dot{x}=\mathrm{d}x/\mathrm{d}t$）の二乗と二回微分（$\ddot{x}=\mathrm{d}^2x/\mathrm{d}t^2$）の条件付き平均を定義する．

$$q(\xi)\equiv\langle\dot{x}^2|x=\xi\rangle/\langle\dot{x}^2\rangle. \quad (22)$$

$$r(\xi)\equiv\langle\ddot{x}|x=\xi\rangle/\langle\dot{x}^2\rangle. \quad (23)$$

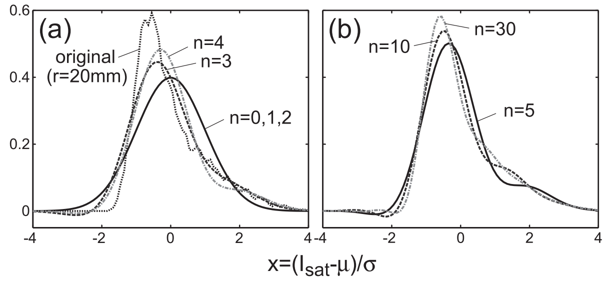

図5　図1に示した信号波形のPDFを直交展開して，式(17)に従って再構成した結果．

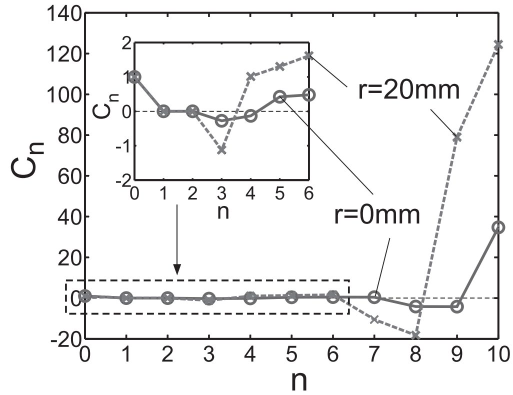

図6　Gram-Charlier 級数展開係数の分布．

条件付き平均法については改めて解説するが，簡略に言えば確率変数がある特定の値を持つことを条件として，統計平均を計算することである．このとき，PDF は $q(\xi)$ と $r(\xi)$ によって一意に以下のように記述される．

$$p(x)=\frac{D}{q(x)}\exp\left[\int_0^x \frac{r(x')}{q(x')}\mathrm{d}x'\right]. \tag{24}$$

ここで，D は $\int_{-\infty}^{+\infty} p(x)dx=1$ の条件を満足する規格化定数である．例えば，$q(x)$ が一定値で $r(x)=-x$ だとすると，$p(x)$ は正規分布となる．

さて，$x(t)$ を流体中で計測される速度変動とする．すると，その時間微分は加速度となるが，エネルギー散逸量 ε[1]に比例すると解釈することも可能である．$\varepsilon\propto(\mathrm{d}u/\mathrm{d}x)^2\propto(\mathrm{d}u/\mathrm{d}t)^2$ と見積もられ[3,5]，瞬時速度 u の条件付き平均が $q(x)$ に他ならない．エネルギー散逸は流体の小さなスケールの運動に起因する物理量である．変動速度のPDF型にこのような具体的な物理量の関連を見出せることは興味深い．また，条件付き平均 $p(x)$ や $q(x)$ の分布自体を調べることも，確率変数の特徴を見出すことにつながる．

図7(a)は乱流境界層[2]中で計測された速度変動の $q(u)$ と $r(u)$ の一例である．壁からの距離（y^+）を変化させた結果を示してある．$q(u)$ はほぼ一定，$r(u)$ は $-2\leq u\leq +2$ の範囲で直線的に変化している．この分布を忠実に用いると，式(24)からPDFを構成できる．PDFは正規分布に近いので，$p(x)$，$q(x)$ の平均値まわりを次の1次関数で近似してみる．

$$q(u)=a_1u+b_1,\qquad r(u)=a_2u+b_2. \tag{25}$$

式(25)を用いて再構成されたPDFを**図8**に示す．$-1\leq u\leq 1$ の範囲で分布は正の方向に少し偏っているが，これは係数 a_1 が負の値を持つことに起因する．**図7**(b)係数 a_1 と a_2 の分布を壁からの距離に対してプロットしたものである．境界層中では，粘性低層[3]，対数領域[4]，後流領域[5]と異なる物理現象が支配する速度分布が形成される．これらの領域とよく対応して係数 a_1，a_2 が分布していることがわかる．特に，対数領域では係数の値が変化せず，したがって $(\partial u/\partial t)^2$ と $(\partial^2 u/\partial t^2)$ が壁からの領域に依存せず，普遍な型をもっていると考えられる．言いかえれば，条件付きエネルギー散逸率が変動速度に依存せずに一定になることと解

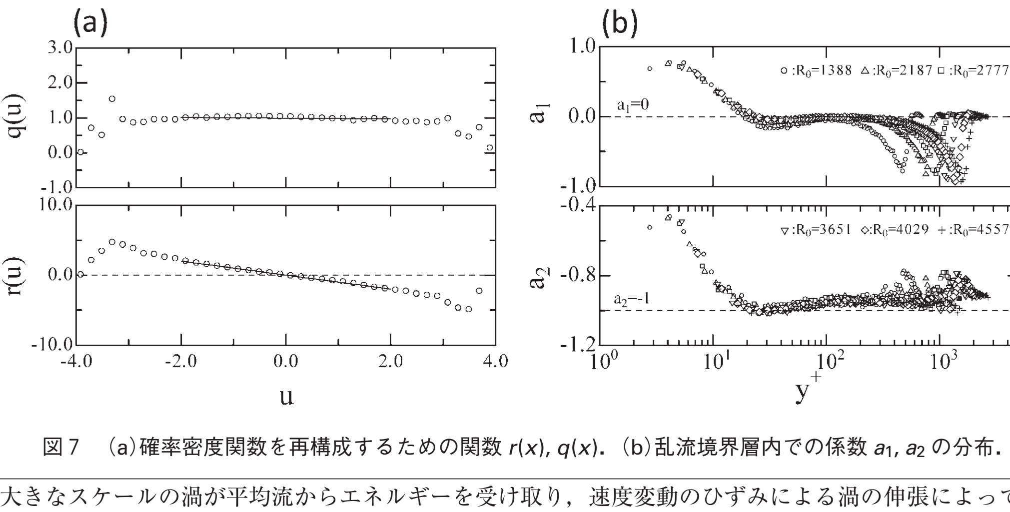

図7　(a)確率密度関数を再構成するための関数 $r(x)$, $q(x)$．(b)乱流境界層内での係数 a_1, a_2 の分布．

1　乱流では大きなスケールの渦が平均流からエネルギーを受け取り，速度変動のひずみによる渦の伸張によって小さなスケールへエネルギーが伝達され，最後に分子粘性によって熱に変換される（エネルギーカスケードと呼ばれる）．乱流エネルギーが分子粘性によって熱に変換されることを粘性散逸，粘性散逸によって熱に変換されるエネルギー量を散逸率と定義する．

2　固体壁面上を流体が流れる場合に形成される，速度が連続的に変化する薄い層．

3　層流底層とも呼ばれ，壁面のごく近傍で分子粘性の作用が卓越する薄い層．壁表面では速度はゼロとなる．

4　粘性底層外層のバッファ層（遷移層）のさらに上層に位置し，平均速度分布に対数則（半理論的に得られた法則）が成立する．乱流境界層の厚さの1/10程度．

5　対数領域の上層に位置し，乱流／非乱流が混在する領域．

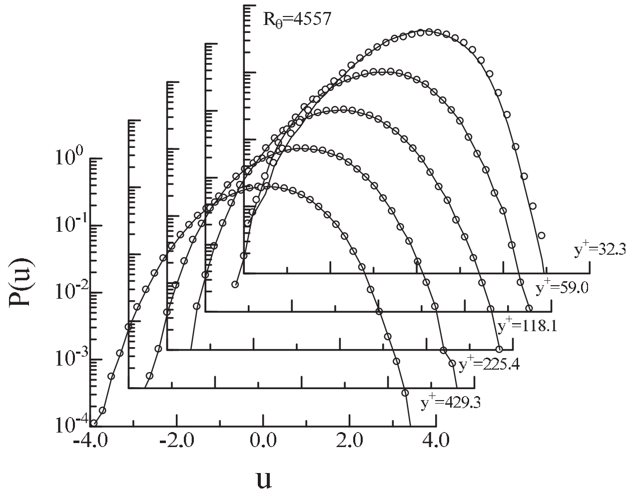

図8　関数 $r(x)$，$q(x)$を用いて再構成された PDF の例（乱流境界層）[15]．白丸は実験点，実線は式(24)をあらわす．

釈される．対数領域においては，大きな散逸率を担う特徴的なスケール変動は存在せず，無次元速度を基準にすれば，空間的に同一の散逸率分布が存在する．対数領域をこのように特徴づけることも可能である．

3.3.4　ダイバージェンス

実験データなどから計算された PDF がどのようなパラメータによって定量化されるかを，いくつかの例を挙げながら説明してきた．本節では，PDF 型の類似度を判断する客観的指標として，カルバックライブラーのダイバージェンス（Kullback-Leibler Divergence 以後，KLD と略す）について紹介したい[17]．KLD は微分情報幾何学で広く利用されており[18]，PDF を族とする関数空間の距離関数になる．二つの離散化された PDF に対して，

$$D(P\|Q)\equiv\sum_{\{\mathbf{s}\}}P(s_i)\log_e(P(s_i)/Q(s_i)),\tag{26}$$

と定義される．ここで，$P(\mathbf{s})$ と $Q(\mathbf{s})$，$\{\mathbf{s}\}=\{s_1,s_2,\cdots\}$ は離散化された PDF である．KLD は非負の値を持ち，二つの PDF 型が同一の場合にはゼロとなり，また，両者の差異が大きいほど大きな値をとる．ただし，距離関数が満たすべき交換則は満たされない（つまり，$D(P\|Q)\neq D(Q\|P)$）．簡単な例としては，$Q(\mathbf{s})$ を正規分布として，$P(\mathbf{s})$ を計測された結果とすると，KLD によって定量的に正規分布からの差異を特徴づけられる．KLD には PDF 型のすべての高次モーメントの情報が含まれているので，より高度な PDF 型の比較をすることができる．

仮に $P(\mathbf{s})$ と $Q(\mathbf{s})$ の差異が小さい場合，即ち $[P(\mathbf{s})-Q(\mathbf{s})]/P(\mathbf{s})\ll 1$，には式(26)は，

$$\begin{aligned}D(P\|Q)&=\sum_{\{\mathbf{s}\}}[Q(s_i)-P(s_i)]+\frac{1}{2}\sum_{\{\mathbf{s}\}}\frac{[Q(s_i)-P(s_i)]^2}{Q(s_i)}+\cdots\\&\simeq\frac{1}{2}\sum_{\{\mathbf{s}\}}\frac{[Q(s_i)-P(s_i)]^2}{Q(s_i)},\end{aligned}\tag{27}$$

となる．KLD をユークリッド空間での距離関数（$E(P\|Q)$）との関連で理解するためには，式(26)の各項を $1/\sqrt{Q(s_i)}$ で重みづけしたものと考えればよい．したがって，

$$E(P\|Q)\equiv\sqrt{\sum_{\{\mathbf{s}\}}[Q(s_i)-P(s_i)]^2},\tag{28}$$

また，KLD と類似の距離関数として Hellinger distance，$H(P\|Q)$ がある．P が十分に Q と類似する場合（$P\simeq Q$）には，次のように KLD との関連をつけることも可能である．KLD の代わりに $H(P\|Q)$ を用いてもよい[19,20]．

$$\begin{aligned}H(P\|Q)&\equiv\sum_{\{\mathbf{s}\}}[\sqrt{Q(s_i)}-\sqrt{P(s_i)}]^2\\&=\sum_{\{\mathbf{s}\}}\frac{[Q(s_i)-P(s_i)]^2}{[\sqrt{Q(s_i)}+\sqrt{P(s_i)}]^2}\\&\simeq\sum_{\{\mathbf{s}\}}\frac{[Q(s_i)-P(s_i)]^2}{4Q(s_i)}\simeq\frac{1}{2}D(P\|Q).\end{aligned}\tag{29}$$

2つの PDF 型の類似度を定量的に判断するために，$K(P\|Q)$，$E(P\|Q)$，$H(P\|Q)$ について簡単に説明をおこなった．比較する PDF 型にもよるが，PDF 型の微妙な差異を識別するためには KLD を用いるのがよい．KLD の計算にあたって，まず PDF を算出する必要がある．互いの掛け算の積分（総和）が必要であるので，bin size を統一し，積分範囲についても定めておかなくてはならない．

図9は境界層中で計測された PDF 型の KLD の分布を示したものである．x，y 軸は壁からの距離を示し，その位置における互いの PDF 型から計算される KLD を z 軸に表示してある[15]．KLD が小さな範囲は PDF 型が類似していることを示しており，空間内での PDF 型の相違を一目で把握することができる．また，プラズマデータの解析例は，文献[21]を参照されたい．

3.4　結合確率密度関数

前章までは確率変数が1つの場合を説明してきたが，本節では2つの場合（プラズマ計測では，密度と温度を同時測定する場合や2点での計測を行う場合）について考える．また，確率変数は n 個の場合に一般化して同様に考えることができる．

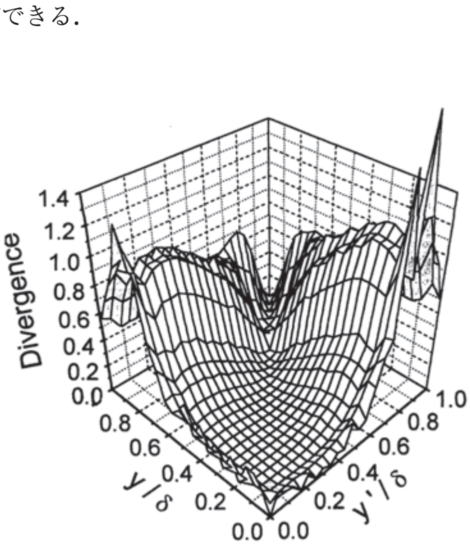

図9　乱流境界層内における KLD の分布．x 軸，y 軸はともに壁からの距離を境界層の厚さで無次元化してある．

3.4.1　定義と具体的な計算法

二つの確率変数xとyがあり，xがある値ξよりも小さな値をとるとき，同時にyもある値ηよりも小さな値をとる確率を結合確率分布関数と呼ぶ．

$$P(\xi,\eta) = \mathrm{prob}[x \le \xi\,;\, y \le \eta]. \tag{30}$$

結合確率分布関数が連続の場合には，一変数の場合と同様に，結合確率密度関数$p(\xi,\eta)$（Joint Probability Density Function；以降，JPDFと略記する）を定義できる．

$$p(\xi,\eta) = \lim_{\Delta\xi,\Delta\eta\to 0}\left[\frac{\mathrm{prob}[x\le\xi; y\le\eta]}{\Delta\xi\Delta\eta}\right]. \tag{31}$$

JPDFは非負の値を持ち（$p(\xi,\eta)\ge 0$），

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(\xi,\eta)\,\mathrm{d}\xi\,\mathrm{d}\eta = 1. \tag{32}$$

が成り立つ．また，結合確率分布関数とは以下の関係式でむすばれる．

$$p(\xi,\eta) = \frac{\partial}{\partial\eta}\left[\frac{\partial P(\xi,\eta)}{\partial\xi}\right]. \tag{33}$$

JPDFは1変数の確率密度関数の情報も含んでいるので，PDFとの関係は，

$$p(\xi) = \int_{-\infty}^{+\infty} p(\xi,\eta')\,\mathrm{d}\eta',\quad p(\eta) = \int_{-\infty}^{+\infty} p(\xi',\eta)\,\mathrm{d}\xi', \tag{34}$$

となる．離散データからJPDFを計算するには，2つの確率変数の標準偏差に基づきbin sizeを定める．2次元平面で，bin sizeを元に区切られた格子状の領域に入る離散データの個数を式(4)にならって個数をカウントすればよい．**図10**には，NAGDIS-IIで観測された電子密度と電子温度とのJPDFを示した．各々平均値と標準偏差値で無次元化してある．電子密度がある値をとったとき，どれくらいの確率である値の電子温度が実現されるかを知ることができる．JPDFの等値面で，y軸の情報を全てx軸に射影したものが$p(\xi)$，x軸の情報を全てy軸に射影したものが$p(\eta)$となる．つまり，ある特定のξ_1に対して，1つのPDF $p(\eta)$の分布が存在するので，JPDFを精度よく計算するためには膨大なデータ数が必要となる．結合確率密度関数が各々の変数のPDFの積として表される場合；$p(\xi,\eta)=p(\xi)p(\eta)$には，確率変数$x$と$y$は互いに独立と定義される．

3.4.2　期待値と相互相関

二つの確率変数xとyを引数とする関数$f(x,y)$の期待値は，JPDFを用いて次のように計算される．

$$E[f(x,y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(\xi,\eta)\,p(\xi,\eta)\,\mathrm{d}\xi\,\mathrm{d}\eta. \tag{35}$$

JPDFの分布を眺めて気づくことは，例えば**図10**(b)では，大きな正の密度には絶対値の大きな電子温度が対応して起こる確率が高く，符号が逆の場合にも同様の傾向があることである．つまり，両者の変動は呼応しており，統計的には正の相関があると言われる．$f(x,y)=xy$の場合には，期待値はこの相関の大きさを定量化したものと解釈することが可能である．JPDFの等値面が原点に対称に分布していれば，$E[f(x,y)]=0$となり，2つの変量は確率的に相関がないと判断される．分布が第二象限と第四象限に偏っていれば，$E[f(x,y)]<0$となり，両者は互いに負の相関を持つと言われる．スペクトルとの関連（第2章）で解説した自己相関関数を相互相関関数と比較していただきたい．

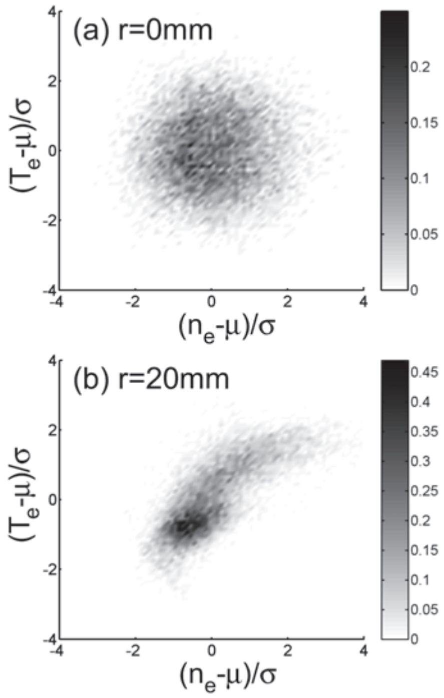

図10　電子密度と電子温度との同時確率密度関数．

3.4.3　確率変数の和と積

二つの確率変数ξ，ηと関数$g(x,y)$が与えられた場合，新たな確率変数$\zeta=g(\xi,\eta)$が定義される．このとき，ξとηのJPDFによって，ζの統計量を計算することを考える．

確率変数ξとηのJPDFが次式に従う場合，それらは結合正規分布と呼ばれる．

$$p(\xi,\eta) = A\exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{(\xi-\mu_1)^2}{\sigma_1^2} - 2r\frac{(\xi-\mu_1)(\eta-\mu_2)}{\sigma_1\sigma_2} + \frac{(\eta-\mu_2)^2}{\sigma_2^2}\right]\right\}, \tag{36}$$

ここで，μ_1，μ_2はξとηの平均値，σ_1，σ_2は標準偏差をあらわす．また，$A=1/\{2\pi\sigma_1\sigma_2\sqrt{1-r^2}\}$，$r$は相互相関係数（$E[((x-\mu_1)(y-\mu_2)]/\sigma_1/\sigma_2$，第2章で説明した）である．二つの確率変数が結合正規分布に従えば，$p(\xi)$および$p(\eta)$はともに正規分布に従う．しかし，逆に$p(\xi)$と$p(\eta)$がともに正規分布に従った場合にも，必ずしも結合正規分布が成り立つとは限らない．この点を注意する必要がある．

関数$g(x,y)$の最も単純な場合として，$g(x,y)=x+y$つまり$\zeta=\xi+\eta$を考える．ζのPDFは以下の手順によって計算される．**図11**を参照して，$\xi+\eta\le\zeta$における左側の斜線部を積分すると，結合確率分布関数

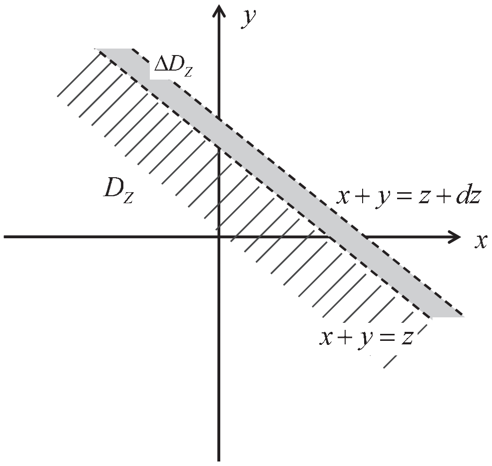

図11　2つの確率変数の積分領域.

$$P(\zeta)=\int_{-\infty}^{+\infty}\int_{-\infty}^{\zeta-\eta}p(\xi',\eta')\mathrm{d}\xi'\mathrm{d}\eta',\tag{37}$$

が得られる．式(33)に従って，$P(\zeta)$ を微分することによって，ζ の確率密度関数が計算される．$z<x+y<z+\mathrm{d}z$ なる領域 ΔD_z は，直線 $x+y=z$ と $x+y=z+\mathrm{d}z$ に囲まれた対角線状の帯域である．この領域の点の座標は$(z-y,y)$であり，微小面積は $\mathrm{d}y\mathrm{d}z$ に等しい．したがって，

$$p_z(\zeta)\mathrm{d}\zeta=\int_{-\infty}^{+\infty}p(\zeta-\eta',\eta')\mathrm{d}\eta'\mathrm{d}\zeta,\tag{38}$$

から直接に計算することもできる．ξ と η が独立の場合には，$p(\xi,\eta)=p_x(\xi)p_y(\eta)$ が成り立つので，上式は

$$p_z(\zeta)=\int_{-\infty}^{+\infty}p_x(\zeta-\eta')p_y(\eta')\mathrm{d}\eta',\tag{39}$$

が得られる．つまり，2つの確率変数が互いに独立ならば，それらの和の密度関数はそれぞれの密度関数の合成積に等しい．

$\zeta=\xi/\eta$ の場合にも，同様の議論が可能である．$z<x/y<z+dz$ なる領域は，直線 $\xi=\eta\zeta$ と直線 $\xi=\eta(\zeta+\mathrm{d}\zeta)$ に囲まれた三角形となる．この領域内の点の座標は(zy,y)であり，微小面積は$|y|\mathrm{d}\eta\mathrm{d}\zeta$ に等しい．したがって，

$$p_z(\zeta)=\int_{-\infty}^{+\infty}|y|p(\zeta'\eta',\eta')\mathrm{d}\eta',\tag{40}$$

が得られる．結合確率密度関数が結合正規分布の場合には，$p_z(\zeta)$ の具体的な型を解析的に導くことができる．

流体乱流でよく計算されるのは，$\zeta=\xi\eta$ の場合である．ξ と η が乱流中の速度変動の場合には，ζ は（瞬時の）レイノルズ応力と呼ばれ，乱れエネルギーの生成などに主要な役割をはたす．x 軸方向の変動速度を u，y 軸方向を v とすると，$\zeta=u\times v$ の分布関数は，

$$P_z(\zeta)=\int_0^{+\infty}\mathrm{d}u\int_{-\infty}^{\zeta/u}\mathrm{d}vp(u,v)+\int_{-\infty}^{0}\mathrm{d}u\int_{\zeta/u}^{\infty}\mathrm{d}vp(u,v),\tag{41}$$

となり，これを微分すると確率密度関数は，

$$p_z(\zeta)=\int_0^{+\infty}\frac{\mathrm{d}u}{u}[p(u,\zeta/u)+p(-u,-\zeta/u)],\tag{42}$$

となる．

図12は乱流境界層中で計測されたPDFの一例である．$\zeta=0$ まわりに鋭いピークがあり，PDFのすそ野が広く延びている特徴がある．ζ のPDFを再構成する方法について考えてみたい．乱流中の変動速度のPDFは，正規分布に近いがそこから少しだけずれる傾向にある．この点を考慮すると，PDF型をある次数までのGram-Charlier級数展開により近似し，それらを用いて結合確率密度関数を算出することが可能である．

$$p(u,v)=\frac{1}{2\pi}\sum_{j+k}C_{jk}\phi^{(j)}(u)\phi^{(k)}(v)\exp\left(-\frac{1}{2}u^2-\frac{1}{2}v^2\right).\tag{43}$$

上式を式(42)に代入すると，

$$p_z(\zeta)=\sum_{j+k=even}\frac{1}{\pi}C_{jk}\int_0^{+\infty}\phi^{(j)}(u)\phi^{(k)}\left(\frac{\zeta}{u}\right)\exp\left[-\frac{1}{2}\left(u^2+\frac{\zeta^2}{u^2}\right)\right]\frac{\mathrm{d}u}{u}.\tag{44}$$

詳しい内容は参考文献[14]にゆずるが，定式は多少の計算の後，第二種の修正ベッセル関数を用いてより簡素な形式に表記される[22]．**図12**に計算の結果を比較のために載せ

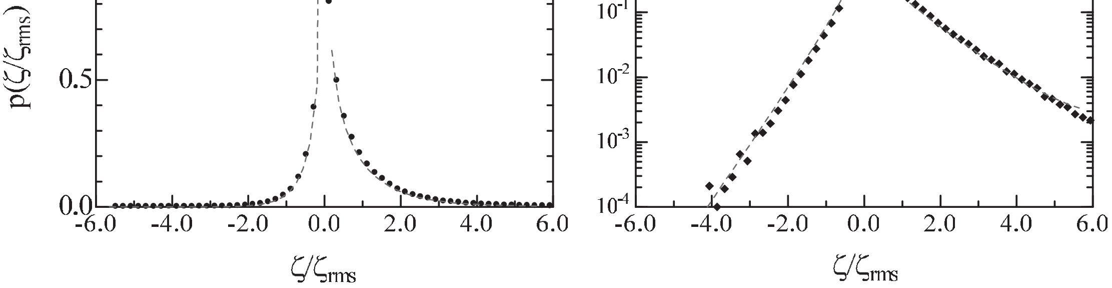

図12　2つの確率変数の積により構成される確率変数のPDF．波線は式(44)より計算される結果.

た．

確率変数が２つの場合にも，１変数の場合と同様にモーメントが定義される．また，１つの確率変数を特定の値 $v = v_*$ に固定して，同時確率分布 $p(u, v_*)$ を調べることもできる．この分布は条件つき確率分布と呼ばれ，式(24)で用いた $q(x)$，$r(x)$ を計算する際に用いられる．

参考文献

［1］伊藤　清：確率論の基礎［新版］（岩波書店，2004）．
［2］J.M. Dewhurst, B. Hnat, N. Ohno, R.O. Dendy, S. Masuzaki, T. Morisaki, and A. Komori, Plasma Phys. Control. Fusion **50**, 095013 (2008).
［3］後藤俊幸：乱流理論の基礎（朝倉書店，1998）．
［4］H. Takayasu, *Fractals in Physical Science* (Manchester University Press, New edition 1991).
［5］木田重雄，柳瀬眞一郎：乱流力学（朝倉書店，1999）．
［6］P.A. Davidson, *Turbulence* (Oxford University Press, 2004).
［7］J.P. Graves, J. Horacek, R.A. Pitts, and K.I. Hopcraft, Plasma Phys. Control. Fusion **47**, L1-L9 (2005).
［8］Athanasios Papoulis, Probability, *Random Variables and Stochasitic Processes,* 中山謙二他訳（東海大学出版会，第１刷，1996）．
［9］G.Y. Antar, G. Counsell, Y. Yu, B. LaBombard, and P. Devynck, Phys. Plasmas **10**, 419 (2003).
［10］V.P. Budaev, S. Masuzaki, T. Morisaki, N. Ohno, N. Asakura, S. Takamura, H. Yamada, and A. Komori, Plasma Fusion Res. **3**, S1019 (2008).
［11］日野幹雄：スペクトル解析（朝倉書店，第17刷，1989）．
［12］N. Ohno, S. Masuzaki, H. Miyoshi, S. Takamura, V.P. Budaev, T. Morisaki, N. Ohyabu, and A. Komori, Contrib. Plasma Phys. **46**, 692 (2006).
［13］O. Barndorff-Nielsen, Proc. R. Soc. Lond. A **368**, 501 (1979).
［14］F. Durst, J. Jovanovic, and Lj. Kanevce, *Turbulent Shear Flows 5* (Springer-Verlag Berlin Heidelberg, 1987) p.197.
［15］Y. Tsuji and B. Lindgren and A.V. Johansson, Fluid Dyn. Res. **37**, 293 (2005).
［16］S.B. Pope and E.C.S. Ching, Phys. Fluids **5**, 1529 (1993).
［17］S. Kullback, *Information Theory and Statistics* (Wiley, London, 1959).
［18］S. Amari, *Lecture Notes in Statistics 28,* (Springer-Verlag,1985).
［19］A.N. Shiryaev, *Probability,* 2nd ed. (Springer-Verlag, New York, 1996).
［20］H. Heyer, *Theory of statistical experiments* (Springer-Verlag, New York, 1982).
［21］V.P. Budaev, Y. Kikuchi, Y. Uesugi and S. Takamura, Nucl. Fusion **44**, S108 (2004).
［22］R.A. Antonia and J.D. Atkinson, J. Fluid Mech. **58**, 581 (1973).