第 53 卷 第 9 期 2004 年 9 月
1000-3290/2004/53(09)/2852-12

物 理 学 报
ACTA PHYSICA SINICA

Vol.53,No.9,September,2004

非平衡统计信息理论

邢修三

(北京理工大学物理系,北京 100081)

(2003 年 6 月 30 日收到;2003 年 8 月 20 日收到修改稿)

阐述了以表述信息演化规律的信息(熵)演化方程为核心的非平衡统计信息理论.推导出了 Shannon 信息(熵)的非线性演化方程,引入了统计物理信息并推导出了它的非线性演化方程.这两种信息(熵)演化方程一致表明:统计信息(熵)密度随时间的变化率是由其在坐标空间(和态变量空间)的漂移、扩散和减损(产生)三者引起的.由此方程出发,给出了统计信息减损率和统计熵产生率的简明公式、漂移信息流和扩散信息流的表达式,证明了非平衡系统内的统计信息减损(或增加)率等于它的统计熵产生(或减少)率、信息扩散与信息减损同时发生,得到了反映传递过程动态特性的动态互信息公式和动态信道容量公式,讨论了 Shannon 信息(熵)及其演化方程和统计物理信息(熵)及其演化方程的同异.

关键词:统计信息(熵)演化方程,统计信息减损率,统计熵产生率,信息(熵)流,信息(熵)扩散,动态互信息,动态信道容量

PACC:0520,0540,8000

1. 引 言

信息理论,虽然人们研究的范围广泛,但是概念明确、数学表述和运算严格、自成体系且能付诸实用的,当前只有 Shannon 建立[1—3]并经 Jaynes 等人[4,5]发展的统计信息理论,即 Shannon 统计信息理论.信息和信息熵,统计信息理论中最基本的概念和量,不仅用作信息量的度量,还用来表示自然和社会各种系统的有序度和无序度[6,7],并广泛用于通信、计算机、控制论、数学、物理学、化学、生物学、经济学、社会学及各种工程科学等领域[5—9].尽管如此,现有 Shannon 统计信息理论发展较为成熟的,仍限于与时空过程无关的静态或平衡态.实际上,无论从信息传递的角度或从动力学系统有序度和无序度的角度来看,信息和信息熵总是随时空过程变化的.近些年,虽然信息动力学[10,11],信息流[12]和信息扩散[13]等概念已有应用,但主要还是定性的,缺乏定量的规律性的论著.近几年,作者[14—17]提出了一个新的非平衡态统计物理基本方程,即 $6N$ 维相空间反常朗之万方程或其等价的刘维扩散方程,以取代现有的刘维方程,进而得到了非平衡熵密度随时空变化的非线性演化方程,预言了熵扩散的存在,给出了熵产生率即熵增加定律的一个简明公式.面对这些非平衡态或动态信息和熵的内容,人们自然会问:在非平衡态或动态系统中,Shannon 信息和信息熵究竟如何随时间、空间和其他态变量有规律地变化?是否遵守什么演化方程?若是这种方程是什么形式?既然已有统计物理熵及其演化方程,是否应存在对应的统计物理信息及其演化方程?它又是什么形式?与 Shannon 信息演化方程有何同异?信息熵产生、信息减损、信息流和信息扩散如何表述?与哪些变量有关?它们与演化方程又有何联系?动态信息传递过程中的互信息公式和信道容量公式是否有所修改?如何修改?它与现有的静态互信息公式和静态信道容量公式有何不同?本文将 Shannon 信息和统计物理信息统一称之为统计信息、信息熵和统计物理熵统一称之为统计熵,以统计信息和统计熵的演化方程为核心来统一解答这些随时空变化的统计信息和统计熵课题,给出了定量的数学表达式.

2. 统计熵演化方程

统计物理熵(简称物理熵),Shannon 信息熵(简称信息熵),都具有统计性,共性显著,本文统一称之为统计熵.它们的演化方程则称之为统计熵演化方程.

2.1. 物理熵演化方程

如上所述,近几年,作者得到了非平衡熵密度随时空变化的非线性演化方程,它在 $6N$ 维、6 维和 3 维相空间的形式如下[15,16]:

$$\frac{\partial S_X}{\partial t} = -\nabla_X \cdot (\dot{X} S_X) + D\nabla_q^2 S_X + \frac{D}{k\rho}[(\nabla_q \ln\rho) S_X - \nabla_q S_X]^2, \qquad (1)$$

$$\frac{\partial S_v}{\partial t} = -\nabla_{q_1} \cdot (C S_v + J_v) + D\nabla_{q_1}^2 S_v + \frac{D}{k}\int \frac{1}{f}[(\nabla_{q_1}\ln f) S_{vp} - \nabla_{q_1} S_{vp}]^2 \mathrm{d}p_1, \qquad (2)$$

$$\frac{\partial S_{vp}}{\partial t} = -\nabla_{q_1} \cdot (v_1 S_{vp} + J_{vp}) + D\nabla_{q_1}^2 S_{vp} + \frac{D}{kf}[(\nabla_{q_1}\ln f) S_{vp} - \nabla_{q_1} S_{vp}]^2, \qquad (3)$$

其中 t 为系统演化的时间,$X = (q,\ p)$ 为 $6N$ 维相空间的状态向量,q 和 p 为一组向量 $q = (q_1,\ q_2,\ \cdots,\ q_N)$ 和 $p = (p_1,\ p_2,\ \cdots,\ p_N)$,$q_i$ 和 p_i 为第 i 个粒子的位置和动量,$v_1 = p_1/m$ 为粒子速度,$x_1 = (q_1,\ p_1)$ 为 6 维相空间的状态向量. $\rho(X,t) = \rho(q,p,t) = \rho(q_1,q_2,\cdots,q_N;p_1,p_2,\cdots,p_N;t)$ 为系综概率密度,$f(x_1,t) = f(q_1,\ p_1,t)$ 为单粒子概率密度. C 为流体平均速度,D 是粒子的自扩散系数. S_X 和 S_{vp} 为 $6N$ 维和 6 维相空间的熵密度,$S_v = \int S_{vp}\mathrm{d}p_1$ 为三维空间的熵密度,J_{vp} 和 $J_v = \int J_{vp}\mathrm{d}p_1$ 为位能在 6 维和 3 维空间贡献的熵流密度,k 为 Bolzmann 常数.

方程(1),(2),(3)的形式相同,它们表明非平衡物理熵密度随时间的变化率(左边)是由其在空间的漂移(右边第一项)、扩散(右边第二项)和产生(右边第三项)三者共同引起的.

正是在这种统计物理熵演化方程的启示下,导致了本文非平衡统计信息理论的提出.

2.2. 信息熵演化方程

由于信息熵是物理熵的推广,而物理熵演化方程的基础来自新的非平衡态统计物理基本方程,为此我们自然会猜想:信息熵不仅应存在类似的演化方程,而其基础亦应是系统的态变量演化方程.

设所讨论的动力学系统,如流体动力学、颗粒(如晶粒、恒星等)演化动力学、信息传递动力学等,其状态可由坐标空间的态变量描述之,如流体的速度和温度、颗粒尺度、传递信息的符号等,设 t 为系统的态变量 a 演化的时间,$\dot{a}$ 为 t 时态变量自身变化的速率,$\dot{x}$ 为态变量在坐标空间的变化速率,$p(a,x,t)\mathrm{d}a\mathrm{d}x$ 为 t 时在坐标 x 和 $x+\mathrm{d}x$ 间找到态变量在 a 和 $a+\mathrm{d}a$ 间的概率. 显然,$p(a,x,t)\mathrm{d}a\mathrm{d}x$ 应满足归一化条件 $\int p(a,x,t)\mathrm{d}a\mathrm{d}t = 1$. 由于系统总会受到内外噪声的干扰,因而 $\dot{a}$ 和 $\dot{x}$ 都由漂移变化速率和涨落变化速率两部分组成. 根据随机理论,描述概率密度 $p(a,x,t)$ 的微分方程应是下述 Fokker-Planck 方程[8]:

$$\frac{\partial p(a,x,t)}{\partial t} = -\frac{\partial}{\partial a}[A(a)p(a,x,t)] - v\frac{\partial p(a,x,t)}{\partial x} + \frac{\partial^2}{\partial a^2}[B(a)p(a,x,t)] + Q\frac{\partial^2 p(a,x,t)}{\partial x^2}. \qquad (4)$$

这就是动力学系统的态变量的概率密度变化的微分方程,或简称态变量演化方程. 其中[18]

$$A(a) = h(a) + g(a)\frac{\partial g(a)}{\partial a}, \quad B(a) = g^2(a),$$

$$g(a) = [D_1 + 2\lambda\sqrt{D_1 D_2}\mu(a) + D_2\mu^2(a)]^{1/2},$$

其中 $h(a)$ 为态变量自身的漂移变化速率,D_1 和 D_2 为其加性噪声和乘性噪声的噪声强度或扩散系数,$\mu(a)$ 为乘性噪声的乘性因子;v 为态变量在坐标空间的漂移变化速率,Q 为其噪声强度或扩散系数,λ 为关联常数,且 $0 \leqslant \lambda \leqslant 1$.

根据信息理论及与非平衡态统计物理熵的对应,动力学系统演化 t 时的动态信息熵可定义为[2,7,14,15]

$$S(t) = -\int p(a,x,t)\log\frac{p(a,x,t)}{p_m(a,x)}\mathrm{d}a\mathrm{d}x + S_m = \int S_{ax}(t)\mathrm{d}a\mathrm{d}x + S_m, \qquad (5)$$

其中 $p_m = p_m(a,x)$ 为信息熵达最大值 S_m 时的概率密度,物理学叫平衡态概率密度,满足

$$\frac{\partial p_m}{\partial t} = -\frac{\partial}{\partial a}(Ap_m) - v\frac{\partial p_m}{\partial t} + \frac{\partial^2}{\partial a^2}(Bp_m) + Q\frac{\partial^2 p_m}{\partial x^2} = 0, \qquad (6)$$

其中

$$S_{ax}(t) = -p(a,x,t)\log\frac{p(a,x,t)}{p_m(a,x)} \qquad (7)$$

叫单位坐标空间和单位态变量空间的信息熵密度. 由于 $\Delta S(t)=S(t)-S_{\rm m}$ 是相对熵,故严格言之, $S_{ax}(t)$ 应叫相对信息熵密度.

将(5)式两边对时间 t 求偏导数,并代入态变量演化方程(4)和(6),则可求得信息熵密度 $S_{ax}(t)$ 的演化方程如下:

$$\begin{aligned}\frac{\partial S_{ax}}{\partial t}=&-\frac{\partial}{\partial a}(AS_{ax})-v\frac{\partial S_{ax}}{\partial x}\\&+\frac{\partial^2}{\partial a^2}(BS_{ax})+Q\frac{\partial^2 S_{ax}}{\partial x^2}\\&+\frac{B}{p}\left[\left(\frac{\partial}{\partial a}\log p\right)S_{ax}-\frac{\partial S_{ax}}{\partial a}\right]^2\\&+\frac{Q}{p}\left[\left(\frac{\partial}{\partial x}\log p\right)S_{ax}-\frac{\partial S_{ax}}{\partial x}\right]^2,\end{aligned}\tag{8}$$

信息熵产生密度

$$\begin{aligned}\sigma&=\sigma_a+\sigma_x\\&=Bp\left(\frac{\partial}{\partial a}\log\frac{p}{p_{\rm m}}\right)^2+Qp\left(\frac{\partial}{\partial x}\log\frac{p}{p_{\rm m}}\right)^2.\end{aligned}\tag{9}$$

将(8)式两边对 a 和对 x 积分,则得单位坐标空间信息熵密度 $S_x(t)$ 和单位态变量空间信息熵密度 $S_a(t)$ 的演化方程各为

$$\begin{aligned}\frac{\partial S_x}{\partial t}=&-v\frac{\partial S_x}{\partial x}+Q\frac{\partial^2 S_x}{\partial x^2}\\&+\int\frac{B}{p}\left[\left(\frac{\partial}{\partial a}\log p\right)S_{ax}-\frac{\partial S_{ax}}{\partial a}\right]^2{\rm d}a\\&+\int\frac{Q}{p}\left[\left(\frac{\partial}{\partial x}\log p\right)S_{ax}-\frac{\partial S_{ax}}{\partial x}\right]^2{\rm d}a,\end{aligned}\tag{10}$$

$$\begin{aligned}\frac{\partial S_a}{\partial t}=&-\frac{\partial}{\partial a}(AS_a)+\frac{\partial^2}{\partial a^2}(BS_a)\\&+\int\frac{B}{p}\left[\left(\frac{\partial}{\partial a}\log p\right)S_{ax}-\frac{\partial S_{ax}}{\partial a}\right]^2{\rm d}x\\&+\int\frac{Q}{p}\left[\left(\frac{\partial}{\partial x}\log p\right)S_{ax}-\frac{\partial S_{ax}}{\partial x}\right]^2{\rm d}x,\end{aligned}\tag{11}$$

其中

$$S_x(t)=\int S_{ax}(t){\rm d}a,\quad S_a(t)=\int S_{ax}(t){\rm d}x.$$

在得到方程(10)和(11)时利用了 I_{ax} 及其导数在 $a\to\pm\infty$ 和 $x\to\pm\infty$ 时都为零的边界条件.

方程(8)(10)(11)就是非平衡系统的三种信息熵密度的演化方程,它指明非平衡信息熵密度随时间的变化率是由其在坐标空间和态变量空间的漂移、扩散和产生三者共同引起的.信息熵扩散和产生同时发生,都来自随机噪声干扰.由于信息熵表示系统的无序度,故演化方程(8),(10),(11)就是描述非平衡系统的无序度密度随时间、坐标空间和态变量空间变化的非线性演化方程.

只要知道了 $A(a)$,$B(a)$,v 和 Q,原则上就可由演化方程(8)(10)(11)解得 $S_{ax}(t)$,$S_x(t)$ 和 $S_a(t)$,但由于(9)(10)(11)式是个复杂的非封闭的非线性偏微分方程,严格的求解是较为困难的.

信息熵密度演化方程(8)和信息密度演化方程(15),不难推广到由 n 个变量(可包括三维空间坐标在内)描述的多维动力学系统,且其数学形式相同.

3. 统计信息演化方程

Shannon 信息,与其信息熵对应;本文引入的统计物理信息(简称物理信息),则与统计物理熵对应. 两种信息都具有统计性,共性显著,本文称之为统计信息.它们的演化方程则称之为统计信息演化方程.

3.1. Shannon 信息演化方程

信息,原仅是通信中信息量的简称,现又用来表示系统的有序度,信息密度则表示系统的有序度密度.在动力学系统中,信息总是随时空和其他态变量变化的,其变化规律应由一个微分方程描述之.

根据信息理论,动力学系统演化 t 时的动态信息可定义为[7]

$$\begin{aligned}I(t)&=S_{\rm m}-S(t)\\&=\int p(a,x,t)\log\frac{p(a,x,t)}{p_{\rm m}(a,x)}{\rm d}a{\rm d}x\\&=\int I_{ax}(t){\rm d}a{\rm d}x,\end{aligned}\tag{12}$$

此式实际上就是 Kullback 信息[2,5],其中

$$I_{ax}(t)=p(a,x,t)\log\frac{p(a,x,t)}{p_{\rm m}(a,x)}\tag{13}$$

为单位坐标空间和单位态变量空间的(相对)信息密度.由(12),(5)式和(13),(7)式得

$$I(t)+S(t)=S_{\rm m},\ I_{ax}(t)+S_{ax}(t)=0,\tag{14}$$

即动力学系统的信息(密度)与信息熵(密度)之和为常数.

将(14)式代入(8)式,则得信息密度 $I_{ax}(t)$ 的演化方程

$$\begin{aligned}\frac{\partial I_{ax}}{\partial t}=&-\frac{\partial}{\partial a}(AI_{ax})-v\frac{\partial I_{ax}}{\partial x}+\frac{\partial^2}{\partial a^2}(BI_{ax})\\&+Q\frac{\partial^2 I_{ax}}{\partial x^2}-\frac{B}{p}\left[\left(\frac{\partial}{\partial a}\log p\right)I_{ax}-\frac{\partial I_{ax}}{\partial a}\right]^2\\&-\frac{Q}{p}\left[\left(\frac{\partial}{\partial x}\log p\right)I_{ax}-\frac{\partial I_{ax}}{\partial x}\right]^2,\end{aligned}\tag{15}$$

信息减损密度

$$\rho = \rho_a + \rho_x = -Bp\left(\frac{\partial}{\partial a}\log\frac{p}{p_{\mathrm{m}}}\right)^2 - Qp\left(\frac{\partial}{\partial x}\log\frac{p}{p_{\mathrm{m}}}\right)^2. \quad (16)$$

将(15)式两边对 a 和 x 积分，则得单位坐标空间信息密度 $I_x(t)$ 和单位态变量空间信息密度 $I_a(t)$ 的演化方程各为

$$\frac{\partial I_x}{\partial t} = -v\frac{\partial I_x}{\partial x} + Q\frac{\partial^2 I_x}{\partial x^2} - \int\frac{B}{p}\left[\left(\frac{\partial}{\partial a}\log p\right)I_{ax} - \frac{\partial I_{ax}}{\partial a}\right]^2\mathrm{d}a - \int\frac{Q}{p}\left[\left(\frac{\partial}{\partial x}\log p\right)I_{ax} - \frac{\partial I_{ax}}{\partial x}\right]^2\mathrm{d}a, \quad (17)$$

$$\frac{\partial I_a}{\partial t} = -\frac{\partial}{\partial a}(AI_a) + \frac{\partial^2}{\partial a^2}(BI_a) - \int\frac{B}{p}\left[\left(\frac{\partial}{\partial a}\log I_{ax}\right) - \frac{\partial I_{ax}}{\partial a}\right]^2\mathrm{d}x - \int\frac{Q}{p}\left[\left(\frac{\partial}{\partial x}\log p\right)I_{ax} - \frac{\partial I_{ax}}{\partial x}\right]^2\mathrm{d}x, \quad (18)$$

其中

$$I_x(t) = \int I_{ax}\mathrm{d}a,\ I_a(t) = \int I_{ax}\mathrm{d}x.$$

方程(15)(17)(18)就是我们推导出的动力学系统的三种信息密度的演化方程. 它表明动态信息密度随时间的变化率是由其在坐标空间和态变量空间的漂移、扩散和减损三者共同引起的. 由于信息密度表示系统的有序度密度，故演化方程(15)(17)(18)就是描述动力学系统的有序度密度随时间、坐标空间和态变量空间变化的非线性演化方程.

若系统只有在空间随机扩散而无漂移变化，例如流体粒子的纯扩散运动，信息符号的纯随机变化，则演化方程(17)就简化为

$$\frac{\partial I_x}{\partial t} = Q\frac{\partial^2 I_x}{\partial x^2} - \frac{Q}{p}\left[\left(\frac{\partial}{\partial x}\log p\right)I_x - \frac{\partial I_x}{\partial x}\right]^2, \quad (19)$$

即信息密度随时间的变化率仅是空间扩散和减损二者共同引起的. 注意，这儿的 $p = p(x,t)$. 由于信息扩散和减损的共同起沅是态变量在空间的随机变化，故两者是否发生是同时的，有信息扩散就有信息减损. 这就使得信息演化方程(19)与质量扩散方程不同，不能变成只有扩散而无减损的单纯信息扩散方程[19]. 原因是质量和能量是守恒的，而信息与信息熵并不守恒.

这儿应指出，由于存在着信息和信息熵的扩散、减损或产生，信息和信息熵演化方程(15)(17)(18)(19)和(8)(10)(11)是时间反演不对称的，反映了这种演化过程的不可逆性.

若系统处于平衡态，则 Shannon 信息和信息熵的演化方程(15)(17)(18)(19)和(8)(10)(11)中的信息(熵)密度不随时间变化且信息(熵)流和信息(熵)减损(产生)都为零. 可见现有的 Shannon 平衡态统计信息理论可看成是其对应的非平衡态统计信息理论的一个与时间过程无关的特殊部分.

3.2. 物理信息演化方程

有了信息熵演化方程(8)(10)(11)和其对应的 Shannon 信息演化方程(15)(17)(18)，自然会想到对应于统计物理熵及其演化方程(1)(2)(3)，亦应存在相应的统计物理信息及其演化方程. 为此，本文从 $6N$ 维、6 维和 3 维相空间的非平衡物理熵密度 $S_{\mathbf{X}}(t)$，$S_{\mathrm{v}}(t)$ 和 $S_{\mathrm{vp}}(t)$ 出发，引入对应的 $6N$ 维、6 维和 3 维相空间的非平衡物理信息密度 $I_{\mathbf{X}}(t)$，$I_{\mathrm{v}}(t)$ 和 $I_{\mathrm{vp}}(t)$，它们同样满足关系式(14)，即

$$\begin{cases} I_{\mathbf{X}}(t) + S_{\mathbf{X}}(t) = 0, \\ I_{\mathrm{v}}(t) + S_{\mathrm{v}}(t) = 0, \\ I_{\mathrm{vp}}(t) + S_{\mathrm{vp}}(t) = 0, \end{cases} \qquad \begin{cases} I_{\mathrm{G}}(t) + S_{\mathrm{G}}(t) = S_{\mathrm{G0}}, \\ I_{\mathrm{B}}(t) + S_{\mathrm{B}}(t) = S_{\mathrm{B0}}. \end{cases} \quad (20)$$

这儿 $S_{\mathrm{G}}(t)$ 和 S_{G0} 为 $6N$ 维相空间的非平衡态和平衡态物理熵(或叫 Gibbs 熵)，$I_{\mathrm{G}}(t)$ 为其对应的 $6N$ 维相空间的非平衡态物理信息；$S_{\mathrm{B}}(t)$ 和 S_{B0} 为 6 维相空间的非平衡态和平衡态物理熵(或叫 Boltzmann 熵)，$I_{\mathrm{B}}(t)$ 为其对应的 6 维相空间的非平衡态物理信息. 根据 $S_{\mathrm{G}}(t)$ 和 $S_{\mathrm{B}}(t)$ 的定义[14—16]，即知 $I_{\mathrm{G}}(t)$ 和 $I_{\mathrm{B}}(t)$ 的定义为

$$\begin{cases} I_{\mathrm{G}}(t) = k\int\rho(\boldsymbol{X},t)\ln\dfrac{\rho(\boldsymbol{X},t)}{\rho_0(\boldsymbol{X})}\mathrm{d}\Gamma = \int I_{\mathbf{X}}(t)\mathrm{d}\Gamma, \\ I_{\mathrm{B}}(t) = k\int f(\boldsymbol{x},t)\ln\dfrac{f(\boldsymbol{x},t)}{f_0(\boldsymbol{x})}\mathrm{d}\boldsymbol{x} = \int I_{\mathrm{vp}}(t)\mathrm{d}\boldsymbol{x}, \\ I_{\mathrm{v}}(t) = \int I_{\mathrm{vp}}(t)\mathrm{d}\boldsymbol{p}. \end{cases} \quad (21)$$

可见，只有 $I_{\mathbf{X}}(t)$ 和 $I_{\mathrm{vp}}(t)$ 是引入的，$I_{\mathrm{G}}(t)$，$I_{\mathrm{B}}(t)$ 和 $I_{\mathrm{V}}(t)$ 是推算出的.

由(20)式可见，非平衡统计物理系统的物理信息(密度)与物理熵(密度)之和为常数.

将(20)式中 $I_{\mathbf{X}}(t)$，$I_{\mathrm{v}}(t)$ 和 $I_{\mathrm{vp}}(t)$ 各取代方程(1)(2)(3)中对应的 $S_{\mathbf{X}}(t)$，$S_{\mathrm{v}}(t)$ 和 $S_{\mathrm{vp}}(t)$，即得 $6N$ 维、6 维和 3 维相空间的统计物理信息密度的非线性演化方程为

$$\frac{\partial I_{\mathrm{X}}}{\partial t}=-\nabla_{\mathrm{X}}\cdot(\dot{X}I_{\mathrm{X}})+D\nabla_q^2 I_{\mathrm{X}}-\frac{D}{k\rho}[(\nabla_q\ln\rho)I_{\mathrm{X}}-\nabla_q I_{\mathrm{X}}]^2, \tag{22}$$

$$\frac{\partial I_{\mathrm{v}}}{\partial t}=-\nabla_{q_1}\cdot(\boldsymbol{C}I_{\mathrm{v}}+\boldsymbol{J}_{\mathrm{v}})+D\nabla_{q_1}^2 I_{\mathrm{v}}-\frac{D}{k}\int\frac{1}{f}[(\nabla_{q_1}\ln f)I_{\mathrm{vp}}-\nabla_{q_1}I_{\mathrm{vp}}]^2\mathrm{d}\boldsymbol{p}_1, \tag{23}$$

$$\frac{\partial I_{\mathrm{vp}}}{\partial t}=-\nabla_{q_1}\cdot(\boldsymbol{v}_1 I_{\mathrm{vp}}+\boldsymbol{J}_{\mathrm{vp}})+D\nabla_{q_1}^2 I_{\mathrm{vp}}-\frac{D}{kf}[(\nabla_{q_1}\ln f)I_{\mathrm{vp}}-\nabla_{q_1}I_{\mathrm{vp}}]^2. \tag{24}$$

方程(22)(23)(24)的形式相同,它表明非平衡物理信息密度随时间的变化率(左边)是由其在坐标空间的漂移(右边第一项)、扩散(右边第二项)和减损(右边第三项)三者共同引起的.由于物理熵(密度)表示统计物理系统的无序度(密度),物理信息(密度)自然就表示统计物理系统的有序度(密度),故演化方程(22)(23)(24)就是描述非平衡统计物理系统的有序度密度随时间和坐标空间变化的非线性演化方程.

需要指出,这儿引入统计物理信息的目的,主要是在数学表述上与统计物理熵的严格对应.在统计热力学系统中,经常谈到熵与信息的关系,如系统的熵增加等于信息减损,熵流出等于信息流入等.若统计物理熵由方程(1)(2)(3)和有关公式表述,那末对应的信息应由哪个公式和方程表述呢?显然,它不是(12)式和方程(15)(17)(18)中的 Shannon 信息,而是(21)式和方程(22)(23)(24)中的物理信息.后面(27)式亦明显给出了这种物理信息与对应的物理熵的变化式.

4. 统计熵产生率和统计信息减损率

物理熵产生率、即熟知的熵增加定律,可否由一个简明公式表述之?这是非平衡统计物理中待解决的一个中心课题.最近,作者[17]给出了物理熵产生率、即物理熵增加定律的一个简明公式,它在其 $6N$ 维和 6 维相空间的形式为

$$\begin{aligned}\frac{\mathrm{d}_i S_{\mathrm{G}}}{\mathrm{d}t}&=P_{\mathrm{G}}=kD\,\overline{(\nabla_q\theta_g)^2},\\ \frac{\mathrm{d}_i S_{\mathrm{B}}}{\mathrm{d}t}&=P_{\mathrm{B}}=kD\,\overline{(\nabla_{q_1}\theta_b)^2},\end{aligned} \tag{25}$$

其中 θ_g 和 θ_b 为非平衡统计物理系统在 $6N$ 维和 6 维相空间的微观状态数密度的离开平衡率.根据演化方程(22)(23)(24)右边的物理信息减损项与演化方程(1)(2)(3)右边的物理熵产生项的表达式相等及导出(25)式的相同程序,可导出 $6N$ 维和 6 维相空间的物理信息减损率、即物理信息减损定律的简明公式为

$$\begin{cases}\dfrac{\mathrm{d}_i I_{\mathrm{G}}}{\mathrm{d}t}=P_{\mathrm{IG}}=-kD\,\overline{(\nabla_q\theta_g)^2},\\[2ex] \dfrac{\mathrm{d}_i I_{\mathrm{B}}}{\mathrm{d}t}=P_{\mathrm{IB}}=-kD\,\overline{(\nabla_{q_1}\theta_b)^2}.\end{cases} \tag{26}$$

由于物理熵增加定律表示统计热力学系统的无序度总是趋向增加,故物理信息减损定律则表示统计热力学系统的有序度总是趋向减损.

(25)(26)式表明,非平衡统计物理系统的宏观物理熵产生和物理信息减损是由其微观状态数密度在坐标空间随机地不均匀离开平衡引起的.关于物理信息减损率 P_{IG}(或 P_{IB})随 θ_g(或 θ_b),$\nabla_q\theta_g$(或$\nabla_{q_1}\theta_b$)及 D 的变化规律,与物理熵产生率 P_{G}(或 P_{B})[17]的相同,亦可参考下面有关 Shannon 信息减损率的讨论.

由(25)(26)式得

$$\begin{cases}\dfrac{\mathrm{d}_i S_{\mathrm{G}}}{\mathrm{d}t}=-\dfrac{\mathrm{d}_i I_{\mathrm{G}}}{\mathrm{d}t},\\[2ex] \dfrac{\mathrm{d}_i S_{\mathrm{B}}}{\mathrm{d}t}=-\dfrac{\mathrm{d}_i I_{\mathrm{B}}}{\mathrm{d}t};\end{cases}$$

或

$$\begin{cases}\dfrac{\mathrm{d}_i}{\mathrm{d}t}=(S_{\mathrm{G}}+I_{\mathrm{G}})=0,\\[2ex] \dfrac{\mathrm{d}_i}{\mathrm{d}t}=(S_{\mathrm{B}}+I_{\mathrm{B}})=0.\end{cases} \tag{27}$$

可见非平衡统计物理系统内部的物理熵与物理信息之和的时间变化率为零,即物理熵产生(或减少)率等于其对应的物理信息减损(或增加)率.

现在给出动力学系统的信息熵产生率及其对应的 Shannon 信息减损率.

首先,我们定义动力学系统的概率密度的离开平衡率或简称系统的离开平衡率为

$$\theta=\log\frac{p}{p_{\mathrm{m}}}\approx\frac{\Delta p}{p_{\mathrm{m}}}, \tag{28}$$

离开平衡率在态变量空间和坐标空间的梯度(或偏微商)

$$\begin{cases}\nabla_a\theta=\nabla_a\log\dfrac{p}{p_{\mathrm{m}}},\\[2ex] \nabla_x\theta=\nabla_x\log\dfrac{p}{p_{\mathrm{m}}}.\end{cases} \tag{29}$$

将(29)式代入(9)式,则得动力学系统的信息熵产生

率的表达式为

$$\begin{aligned}\frac{\mathrm{d}_i S}{\mathrm{d}t} &= \int (\sigma_a + \sigma_x)\,\mathrm{d}a\,\mathrm{d}x \\ &= \int p\left[B\left(\frac{\partial\theta}{\partial a}\right)^2 + Q\left(\frac{\partial\theta}{\partial x}\right)^2\right]\mathrm{d}a\,\mathrm{d}x \\ &= \overline{B\left(\frac{\partial\theta}{\partial a}\right)^2} + \overline{Q\left(\frac{\partial\theta}{\partial x}\right)^2} \geqslant 0.\end{aligned} \tag{30}$$

这正是信息熵产生率的简明公式,亦即是信息熵增加定律的简明公式.信息熵增加定律,又可名广义熵增加定律,与物理熵增加定律相比,它不仅可用于自然科学,亦可用于社会科学.它表示自然和社会动力学系统的信息熵总是趋向增加,即系统自发向信息熵增加的方向演化.由于信息熵表示系统的无序度,故信息熵增加定律表示自然和社会动力学系统的无序度总是趋向增加.

同样,将(28)(29)式代入(16)式,则得动力学系统的 Shannon 信息减损率的表达式为

$$\begin{aligned}\frac{\mathrm{d}_i I}{\mathrm{d}t} &= -\int p\left[B\left(\frac{\partial\theta}{\partial a}\right)^2 + Q\left(\frac{\partial\theta}{\partial x}\right)^2\right]\mathrm{d}a\,\mathrm{d}x \\ &= -\left[\overline{B\left(\frac{\partial\theta}{\partial a}\right)^2} + \overline{Q\left(\frac{\partial\theta}{\partial x}\right)^2}\right] \leqslant 0.\end{aligned} \tag{31}$$

这正是 Shannon 信息减损率的简明公式,与信息熵增加定律相对应,此式亦可看成 Shannon 信息减损定律的简明公式.它表明自然和社会动力学系统的信息总是趋向减损,即系统自发向信息减损的方向演化.由于信息表示系统的有序度,故 Shannon 信息减损定律表示自然和社会动力学系统的有序度总是趋向减损.

(30)(31)式表明,信息熵产生率和信息减损率两者都等于扩散系数或噪声强度与离开平衡率梯度平方的乘积之平均值,指明动力学系统的信息熵产生和信息减损是由其概率密度在坐标空间和态变量空间随机地不均匀离开平衡引起的.可见,复杂动力学系统,只要它处于非平衡($\theta \neq 0$)非均匀($\frac{\partial\theta}{\partial a} \neq 0$, $\frac{\partial\theta}{\partial x} \neq 0$)态且具有随机扩散运动($B \neq 0$, $Q \neq 0$),它的信息熵总是在产生,信息总是在减损.反之,当系统或处于平衡态($\theta = 0$)或虽处于非平衡态但却是均匀的$\left(\frac{\partial\theta}{\partial a} = 0, \frac{\partial\theta}{\partial x} = 0\right)$,或只具有确定性运动而无随机性运动($B = 0$, $Q = 0$),它都没有信息熵产生和信息减损.

由(30)(31)式可以看出,动力学系统的信息熵产生则信息减损,信息熵减少则信息增加,两者同时存在,数值相等.精确地说,信息熵产生(或减少)率等于信息减损(或增加)率,即动力学系统内部的信息熵与信息之和随时间的变化率为零.

$$\frac{\mathrm{d}_i S}{\mathrm{d}t} = -\frac{\mathrm{d}_i I}{\mathrm{d}t} \quad 或 \quad \frac{\mathrm{d}_i}{\mathrm{d}t}(S + I) = 0. \tag{32}$$

从信息熵产生定律和信息减损定律出发,可以断定:任何自然和社会动力学系统都不可能纯洁无瑕,铁板一块,其内部总存在无序子系统,例如晶体内有缺陷,健康生物体内有死亡细胞,社会上有犯罪分子等.

5. 信息流和信息扩散

从 Shannon 信息演化方程(15)或(17)(18)可以求得,态变量空间和坐标空间的信息流由漂移信息流和扩散信息流二者相加而成.

态变量空间的信息流密度为

$$J_a^{ax} = J_{ta}^{ax} + J_{da}^{ax} = AI_{ax} - \frac{\partial}{\partial a}(BI_{ax}) \tag{33a}$$

或

$$J_a = J_{ta} + J_{da} = AI_a - \frac{\partial}{\partial a}(BI_a). \tag{33b}$$

这种信息流只引起系统内部的信息密度变化,与环境的信息交换无关.

坐标空间的信息流密度为

$$J_x^{ax} = J_{tx}^{ax} + J_{dx}^{ax} = vI_{ax} - Q\frac{\partial I_{ax}}{\partial x} \tag{34a}$$

或

$$J_x = J_{tx} + J_{dx} = vI_x - Q\frac{\partial I_x}{\partial x}. \tag{34b}$$

这种信息流无论在通信过程中或研究开放系统的有序度变化时,都有其实际意义.这是因为系统与环境的信息交换,包括一个系统与另一个系统的信息交换,生物内部各部分间的信息交换,社会内部各成员间的信息交换,都是通过坐标空间信息流的来回传递而实现的.

坐标空间的漂移信息流密度为

$$J_{tx}^{ax} = vI_{ax} \tag{35a}$$

或

$$J_{tx} = vI_x. \tag{35b}$$

它等于信息在空间的漂移速率 v 和信息密度 I_{ax} 或 I_x 之乘积.如前所述,v 实际上就是态变量在空间的漂移速率.在通信中,当人用眼接受信息时,则 v 是光速;用耳接受信息时,则 v 是声速.

坐标空间的扩散信息流密度为

$$J_{dx}^{ax} = - Q \frac{\partial I}{\partial x} \tag{36a}$$

或

$$J_{dx} = - Q \frac{\partial I_x}{\partial x}. \tag{36b}$$

它等于信息在空间的扩散系数 Q 与信息密度负梯度 $-\frac{\partial I_{ax}}{\partial x}$ 或 $-\frac{\partial I_x}{\partial x}$ 之积.换言之,信息扩散总是从高密度区向低密度区的方向自发进行的.在三维空间,这种扩散方向可以是各向同性的、四面八方的,因而与漂移流的定向性有明显区别.小道消息(谣言)是典型的信息扩散.动物用气味通讯,实际上就是信息扩散通信.互连网中的信息扩散[13],由于其重要性,已有专题讨论.正因为信息具有自发扩散性,各国政治、军事和金融的保密性都受到高度重视.

信息和信息熵扩散的结果,将使信息和信息熵的密度均匀化,从而导致系统趋向平衡.

由于存在着累积减损,信息流密度在空间传递途中会减小.若系统是非平衡态的,则由(34b)式可求出 t 时流经 x 处的信息流密度

$$J_x(t) = \int\left\{\left[vp(a,x,t) - Q\frac{\partial p(a,x,t)}{\partial x}\right] \times \ln\frac{p(a,x,t)}{p_{\mathrm{m}}(a,x)} - Q\left[\frac{\partial p(a,x,t)}{\partial x} - \frac{p(a,x,t)}{p_{\mathrm{m}}(a,x)}\frac{\partial p_{\mathrm{m}}(a,x)}{\partial x}\right]\right\}\mathrm{d}a, \tag{37}$$

可见空间信息流密度是时空的函数.由此不难求出 $t=0$ 时流经 $x=0$ 处与 $t=\tau$ 时流经 $x=l$ 处两者信息流密度之差,即非平衡传递过程中信息流密度随传递距离增加而减小的变化式

$$J_0(0) - J_l(\tau) = P_l(\tau), \tag{38}$$

其中 $P_l(\tau)$ 为信息流密度在传递途中的累积减损,它可由(37)式计算出.

若系统处于定态,由方程(17) $\frac{\partial I_x}{\partial t} = -\frac{\partial j'_x}{\partial x} = 0$ (这儿 $J'_x \neq J_x$),代入(34b)式,则得信息流密度由于传递途中的累积减损而随传递长度增加而减小的变化式

$$J_x = J_0 - \int_0^x \mathrm{d}x\int_{-\infty}^{\infty}\mathrm{d}a\left\{\frac{B}{p_{st}}\left[\left(\frac{\partial}{\partial a}\log p_{st}\right)I_{ax} - \frac{\partial I_{ax}}{\partial a}\right]^2 + \frac{Q}{p_{st}}\left[\left(\frac{\partial}{\partial x}\log p_{st}\right)I_{st} - \frac{\partial I_{ax}}{\partial x}\right]^2\right\}$$

$$= J_0 - P_x, \tag{39}$$

其中 J_0 和 J_x 为流经 $x=0$ 和 $x=x$ 两处的信息流密度,P_x 为累积减损,因是定态,故都与时间无关.

关于物理信息流,同样由漂移信息流和扩散信息流二者组成,可由物理信息密度演化方程(22)(23)(24)中求得,正如物理熵流可由物理熵密度演化方程(1)(2)(3)中求得[14,15]一样,这儿从略.

作为一个例题,我们来给出布朗运动的漂移扩散传递的信息流密度随时空的变化式.设系统的态变量自身作布朗运动,同时在坐标空间漂移扩散传递,如在空间高斯信道中传递的动态信号,流动流体中的布朗粒子运动,就是这方面的例子.在这种传递过程中,系统的信息熵和信息都将发生变化.根据方程(4),描述这种系统的 Fokker-Planck 方程为

$$\frac{\partial p}{\partial t} = \beta\frac{\partial}{\partial a}(ap) - v\frac{\partial p}{\partial x} + D\frac{\partial^2 p}{\partial a^2} + Q\frac{\partial^2 p}{\partial x^2}, \tag{40}$$

其中 β 为阻力系数.设方程(40)的起始条件为

$$p(a,x,t=0) = (\pi\varepsilon_0)^{-1/2}\exp\left[-\frac{(a-a_0)^2}{\varepsilon_0}\right] \times 2(\pi\varphi)^{-1/2}\exp\left(-\frac{x^2}{\varphi}\right), \tag{41}$$

则可得方程(40)的含时解为[20]

$$p(a,x,t) = (\pi\varepsilon_t)^{-1/2}\exp\left[-\frac{[a-b(t)]^2}{\varepsilon_t}\right] \times 2(\pi\varphi + 4\pi Qt)^{-1/2} \times \exp\left[-\frac{(x-vt)^2}{(\varphi+4Qt)}\right]$$

$$= p(a,t)p(x,t), \tag{42}$$

其中

$$\varepsilon_t = \varepsilon_0 e^{-2\beta t} + \varepsilon_m(1-e^{-2\beta t}),$$

$$b(t) = b_0 e^{-\beta t}, \quad \varepsilon_{\mathrm{m}} = \frac{2D}{\beta}.$$

方程(40)的平衡态解为

$$p_{\mathrm{m}}(a,x) = p_{\mathrm{m}}(a)p_{\mathrm{m}}(x),$$

$$\begin{cases} p_{\mathrm{m}}(a) = (\pi\varepsilon_{\mathrm{m}})^{-1/2}\exp\left(-\frac{a^2}{\varepsilon_{\mathrm{m}}}\right), \\ p_{\mathrm{m}}(x) = \text{常数}. \end{cases} \tag{43}$$

将(42)(43)式代入(37)式得该系统于 t 时流经 x 处的空间信息流密度为

$$J_x(t) = vI_x(t) - Q\frac{\partial I_x}{\partial x}$$

$$= \left\{\left[v + \frac{2Q(x-vt)}{\varphi+4Qt}\right]\left[\frac{b_t^2}{\varepsilon_{\mathrm{m}}} + \frac{1}{2}\left(\frac{\varepsilon_t}{\varepsilon_{\mathrm{m}}} - \ln\frac{e\varepsilon_t}{\varepsilon_{\mathrm{m}}}\right) + \frac{1}{2}\ln\frac{(\varphi+4Qt_f)}{\varphi+4Qt} - \frac{(x-vt)^2}{\varphi+4Qt}\right]\right.$$

$$+\frac{2Q(x-vt)}{\varphi+4Qt}\Big\}p(x,t). \tag{44}$$

系统于 $t=0$ 时在 $x=0$ 处的信息流密度 $J_0(0)$ 与 $t=\tau$ 时传递至 $x=l$ 处的信息密度 $J_l(\tau)$ 之差为

$$\begin{aligned}J_0(0)-J_l(\tau)&=P_l(\tau)\\&=v\Big[\frac{p_{00}b_0^2-p_{l\tau}b_\tau^2}{\varepsilon_m}\\&\quad+\frac{1}{2}\Big(\frac{p_{00}\varepsilon_0-p_{l\tau}\varepsilon_\tau}{\varepsilon_m}\\&\quad-p_{00}\ln\frac{e\varepsilon_0}{\varepsilon_m}+p_{l\tau}\ln\frac{e\varepsilon_\tau}{\varepsilon_m}\Big)\\&\quad+\frac{p_{00}}{2}\ln\frac{(\varphi+4Qt_f)}{\varphi}\\&\quad-\frac{p_{l\tau}}{2}\ln\frac{(\varphi+4Qt_f)}{\varphi+4Q\tau}\Big]>0,\end{aligned}\tag{45}$$

其中 $\tau=l/v$，$p_{00}\equiv p(x=0,t=0)=2(\pi\varphi)^{-1/2}$，$p_{l\tau}\equiv p(x=l,t=\tau)=2(\pi\varphi+4\pi Q\tau)^{-1/2}$，$P_l(\tau)$ 为该系统的信息流密度在传递过程中的累积减损.(45)式正是由(40)式表述的特殊非平衡系统在传递过程中信息流密度的变化式，其中 $J_l(\tau)$ 和 $P(\tau)$ 随 τ 和 Q 的变化及物理意义，留待后面的(56)式一并讨论.

6. 动态互信息和动态信道容量

Shannon 信息传递理论的一个重要课题是给出可用于定量计算信宿在输出端收到信沅从输入端通过信道传递的信息量的公式，这就是现有的静态互信息公式[1-3]：

$$\begin{aligned}I(A;B)&=H(A)-H(A\mid B)\\&=H(B)-H(B\mid A),\end{aligned}\tag{46}$$

其中 A 是从输入端输入信道的符号，B 是信宿在输出端收到的符号. $H(A)$ 是信沅熵，$H(B)$ 是信宿熵，$H(A\mid B)$ 和 $H(B\mid A)$ 都是由噪声对符号的干扰引起的噪声熵.这儿要指出的是，信息传递就是信息流从时空坐标中的一点传递至另一点，是个随时空变化的动态过程.若符号的传递速度是 v，则它在长度为 l 的信道中传递所需的时间 $\tau=l/v$.为此需标明符号在信道中的时空坐标，如令 $A_x(t)$ 和 $B_x(t)$ 为 t 时在信道 x 处的符号 A 和 B.(46)式各项中的符号所以不显含时空坐标，是因现有的静态通信模型，实质上相当于一个点模型，没有时空过程，即符号从输入端传递至输出端不需任何时间，两端相当于处在同一个点，噪声沅当然亦集中于此点.这样，(46)式各项中的符号都是同一时间同一点的符号.以互信息而言，则应是

$$\begin{aligned}I(A;B)&=I(A_0(0);B_0(0))\\&=I(A_l(\tau);B_l(\tau)),\end{aligned}$$

它的精确定义是：信宿于 $t=\tau(=0)$ 时在输出端 $x=l(=0)$ 处从符号 $B_l(\tau)(=B_0(0))$ 中提取该地该时由信沅传递来的符号 $A_l(\tau)(=A_0(0))$ 的信息量.现在的问题是，当符号在信道中的传递时间足够长且传递途中存在噪声时，如何计算信宿于 $t=\tau$ 时在输出端 $x=l$ 处从符号 $B_l(\tau)$ 中提取信沅于 $t=0$ 时由输入端 $x=0$ 处发送的符号 $A_0(0)$ 的信息量？这就是动态互信息 $I(A_0(0);B_l(\tau))$.

为了求出动态互信息公式，将静态互信息公式(46)写成

$$\begin{aligned}I(A_l(\tau);B_l(\tau))&=H(A_l(\tau))-H(A_l(\tau)\mid B_l(\tau))\\&=H(B_l(\tau))-H(B_l(\tau)\mid A_l(\tau)),\end{aligned}\tag{46a}$$

与(46)式相比，这个公式各项都显含了传递时间 τ 和输出端 l，其精确定义已如上所述.由于动态信息在传递途中存在着累积减损，信沅在输入端的信息 $H(A_0(0))$ 经由信息流传递到输出端时已减损成 $H(A_l(\tau))$，两者的关系式为

$$H(A_l(\tau))=H(A_0(0))-H(Q_l(\tau)),\tag{47}$$

其中 $H(Q_l(\tau))$ 为信息从输入端传递至输出端在信道途中的累积减损.将(47)式代入(46a)式即得动态互信息公式

$$\begin{aligned}I(A_0(0);B_l(\tau))&=H(A_0(0))\\&\quad-H(A_0(0)\mid B_l(\tau)),\end{aligned}\tag{48}$$

其中

$$\begin{aligned}H(A_0(0)\mid B_l(\tau))&=H(Q_l(\tau))\\&\quad+H(A_l(\tau)\mid B_l(\tau)).\end{aligned}\tag{49}$$

(48)式虽由(46a)式变来，但其物理意义已不同，它所给出的互信息正是信宿于 $t=\tau$ 时在输出端 $x=l$ 处从符号 $B_l(\tau)$ 中提取信沅于 $t=0$ 时由输入端 $x=0$ 处发送的符号 $A_0(0)$ 的信息量.我们所以把(48)式叫动态互信息，就是因其与(46)(46a)式不同，它包含了信息在时空传递过程中的动态减损.

为了使(48)式可用于实际计算，需知信息在传递途中的累积减损 $H(Q_l(\tau))$，将信息流密度在传递途中的变化式(38)或(39)代入(46)式，不难求得

$$H(Q_l(\tau))=\frac{P_l(\tau)}{J_0(0)}H(A_0(0)),$$

$$H(A_l(\tau)) = \left[1 - \frac{P_l(\tau)}{J_0(0)}\right] H(A_0(0))$$
$$= \frac{J_l(\tau)}{J_0(0)} H(A_0(0)). \tag{50}$$

再将(49)(50)式代入(48)式,则动态互信息公式变为

$$I(A_0(0);B_l(\tau)) = \frac{J_l(\tau)}{J_0(0)} H(A_0(0)) - H(A_l(\tau) \mid B_l(\tau)). \tag{51}$$

由于(38)式和(39)式分别适用于非平衡态和定态,故(51)式既可表示非平衡态又可表示定态的动态互信息.

信息传递理论中另一个重要课题是信道所能传递的最大信息量,这就是信道容量.它与互信息密切相关.在现有 Shannon 信息理论中,静态信道容量定义为[1—3]

$$C = \max_{p(a)} I(A;B), \tag{52}$$

其中 $I(A;B)$就是(46)式的静态互信息公式.随着动态互信息公式(48)(51)的出现,自然就有其相应的动态信道容量公式.与(52)式相对应,它应定义为

$$C_d = \max_{p(a_0)} I(A_0(0);B_l(\tau))$$
$$= \max_{p(a_0)}\left[\frac{J_l(\tau)}{J_0(0)} H(A_0(0)) - H(A_l(\tau) \mid B_l(\tau))\right], \tag{53}$$

其中 $p(a_0)$表示于 $t=0$ 时由输入端 $x=0$ 处发送符号 $A_0(0)$的概率密度分布.(53)式就是将(52)式中的静态互信息公式(46)由动态互信息公式(51)取代后变成的.

由(48)(49)(51)(53)式可见,动态互信息公式和动态信道容量公式中噪声引起的信息减损应由两部分组成:其一是(46)(46a)(52)式中那种在输出端接收时的瞬时信息减损 $H(A_l(\tau) \mid B_l(\tau))$,其二则是在信道传递过程中的累积信息减损 $H(Q_l(\tau))$.何种条件下 $H(Q_l(\tau))$可以略去?由(50)式特别由(45)式及后面的(56)式可见,仅当传递时间 $\tau = l/v \to 0$ 即信道长度 l 与信号传递速度 v 之比趋于零,累积信息减损 $H(Q_l(\tau))$才可略去.这种条件下,输入端的信息传递至输出端时几乎无变化,(48)(51)(53)式就变成(46)(46a)(52)式,这亦正是静态互信息公式(46)(46a)和静态信道容量公式(52)适用的条件.反之,若信道长度 l 很长或传递速度 v 很小或信道噪声强度 Q 很大时,累积信息减损 $H(Q_l(\tau))$则不可略去.这种条件下,表达信宿在输出端收到输入端的信息量的互信息公式和信道容量公式应由静态(46)(46a)(52)式改为动态(48)(51)(53)式.海洋水下通信,恒星间的通信,前者因声速较低,后者因距离特大,且两者途中都存在较强的噪声,因而动态互信息公式就有其实际意义.

应该指出,为了与现有的静态信息理论[1—3]完全对应,动力学系统演化 t 时的动态信息熵(5)式和动态信息(12)式应定义为

$$S(t) = -\int p(a,x,t)\log p(a,x,t)\,\mathrm{d}a\,\mathrm{d}x$$
$$= \int S_{ax}(t)\,\mathrm{d}a\,\mathrm{d}x, \tag{5a}$$
$$I(t) = S_{\mathrm{m}} - S(t)$$
$$= S_{\mathrm{m}} + \int p(a,x,t)\log p(a,x,t)\,\mathrm{d}a\,\mathrm{d}x$$
$$= S_{\mathrm{m}} + \int I_{ax}(t)\,\mathrm{d}a\,\mathrm{d}x. \tag{12a}$$

(5a)(12a)与(5)(12)式的形式虽稍不同,但两者同样通用.更重要的是,当 A,B,v 和 Q 近似常数时(信息传递过程中的 v 和 Q 就可看成常数),(5a)式中的信息熵密度 $S_{ax}(t)$和(12a)式中的信息密度 $I_{ax}(t)$所遵守的演化方程与(8)和(15)式相同,差别是与(9)和(16)式的信息熵产生率密度和信息减损率密度在这儿稍有变化.

与(37)式相对应,若系统处于非平衡态,则于 t 时流经 x 处的信息流密度在这儿为

$$J_x(t) = \int\left\{\left[vp(a,x,t) - Q\frac{\partial p(a,x,t)}{\partial x}\right] \times \ln p(a,x,t) - Q\frac{\partial p(a,x,t)}{\partial x}\right\}\mathrm{d}a. \tag{37a}$$

作为例题,本节给出高斯信道的动态互信息和动态信道容量.在 Shannon 的静态信息论中,高斯信道是个典型的例子,它的信号和噪声都是高斯分布的,其互信息公式和信道容量公式[1—3]为

$$H(A;B) = \frac{1}{2}\ln\left(1 + \frac{\sigma_n^2}{\sigma_n^2}\right), \tag{54}$$

$$C = \frac{1}{2}\ln\left(1 + \frac{\sigma_a^2}{\sigma_n^2}\right), \tag{55}$$

式中 σ_a^2 和 σ_n^2 为输入信号和噪声的方差,即两者的平均功率.(55)式就是熟知的 Shannon 信道容量公式.

现在来给出高斯信道中非平衡传递的信号的动态互信息公式和动态信道容量公式.按前面所述,信息传递是通过信息流实现的,而信息流在信道传递

途中由于存在累积减损是要变小的.根据(37)(44)式,高斯信道中非平衡传递的信号,它在 $t=\tau$ 时传递至输出端 $x=l$ 处的信息流密度 $J_l(\tau)$ 与 $t=0$ 时从输入端 $x=0$ 处发出的信息流密度 $J_0(0)$ 之比

$$\delta = \frac{J_l(\tau)}{J_0(0)} = \frac{\left[\frac{b_\tau^2}{\varepsilon_m} + \frac{1}{2}\left(\frac{\varepsilon_\tau}{\varepsilon_m} - \ln\frac{e\varepsilon_\tau}{\varepsilon_m}\right) + \frac{1}{2}\ln\frac{\varphi + 4Qt_f}{\varphi + 4Q\tau}\right](\pi\varphi)^{1/2}}{\left[\frac{b_0^2}{\varepsilon_m} + \frac{1}{2}\left(\frac{\varepsilon_0}{\varepsilon_m} - \ln\frac{e\varepsilon_0}{\varepsilon_m}\right) + \frac{1}{2}\ln\frac{\varphi + 4Qt_f}{\varphi}\right](\pi\varphi + 4\pi Q\tau)^{1/2}}, \tag{56}$$

根据此式和(50)式,在高斯信道中,若 $t=0$ 时从输入端 $x=0$ 发出的信息为

$$H(A_0(0)) = \frac{1}{2}\ln(2\pi e\sigma_a^2), \tag{57}$$

则它于 $t=\tau$ 时传递到输出端 $x=l$ 处应变为

$$\begin{aligned} H(A_l(\tau)) &= \delta H(A_0(0)) \\ &= \frac{\delta}{2}\ln(2\pi e\sigma_a^2) \\ &= \frac{1}{2}\ln(2\pi e\sigma_{a_l}^2), \end{aligned} \tag{58}$$

其中

$$\sigma_{a_l}^2 = (2\pi e)^{\delta-1}\sigma_a^{2\delta}. \tag{59}$$

(58)式仍是高斯分布的信号,只是其平均功率由 σ_a^2 变小为 $\sigma_{a_l}^2$.将它代入(51)(53)式并利用(54)(55)式的算法,即得高斯信道的动态互信息公式和动态信道容量公式

$$I(A_0(0);B_l(\tau)) = \frac{1}{2}\ln\left[1 + \frac{(2\pi e)^{\delta-1}\sigma_a^{2\delta}}{\sigma_n^2}\right], \tag{60}$$

$$C_d = \frac{1}{2}\ln\left[1 + \frac{(2\pi e)^{\delta-1}\sigma_a^{2\delta}}{\sigma_n^2}\right]. \tag{61}$$

从(56)式可见,δ 与传递时间 $\tau=l/v$ 即信道长度 l、传递速度 v 及空间噪声强度 Q 有关.当 $\tau=l/v\to 0$ 即 $l\to 0$ 或 $v\to\infty$,则 $\delta=1$,$J_l(\tau)=J_0(0)$.这种情况下,信息在信道途中累积减损 $H(Q_l(\tau))=0$,$\sigma_{a_l}^2=\sigma_a^2$,动态互信息公式和动态信道容量公式(60)(61)还原为静态互信息公式和静态信道容量公式(54)(55).实际上,$\tau=l/v\to 0$ 物理上意味着输入端和输出端两点缩短为一点,噪声当然亦集中于这一点,这就是现有的静态通信模型.当 $\tau=l/v\to\infty$ 即 $l\to\infty$ 或 $v\to 0$,或 $Q\to\infty$,则 $\delta=0$,$J_l(\tau)=0$.这种情况下,$H(Q_l(\tau))=H(A_0(0))$,$H(A_l(\tau))=0$,信沅从输入端发出符号 $A_0(0)$ 的信息全部减损于传递途中,在输出端收不到有关它的任何信息.在其他一般情况下,即当 $0<\tau<\infty$ 且 $0<Q<\infty$,则 $0<\delta<1$,$\sigma_{a_l}^2<\sigma_a^2$,高斯动态互信息和动态信道容量都小于其静态互信息和静态信道容量.这些结果就是高斯信道的动态互信息和动态信道容量对其静态互信息和静态信道容量的修正.

本文上述结果都是基于动态信息熵(5)式和动态信息(12)式.若从动态信息(12a)出发,将(42)代入(37a)式,即可求得高斯信道中由于传递途中存在累积减损于终态和始态时空两点的信息流密度之比

$$\begin{aligned} \delta &= \frac{J_l(\tau)}{J_0(0)} \\ &= \frac{(\pi\varphi)^{1/2}\ln[(\pi e\varepsilon_\tau)(\pi\varphi + 4\pi Q\tau)]}{(\pi\varphi + 4\pi Q\tau)^{1/2}\ln[(\pi e\varepsilon_0)(\pi\varphi)]}. \end{aligned} \tag{56a}$$

(56a)和(56)式的物理意义是相同的,两者的差别来自其计算的出发点稍有不同.若用(56a)式代替(56)式来计算高斯信道的动态互信息和动态信道容量(60)(61)式,定量的数值虽稍有差别,定性的结果与上述讨论的完全相同.

7. Shannon 信息(熵)和物理信息(熵)的同异

Shannon 信息和物理信息,信息熵和物理熵,虽是不同领域的统计信息和统计熵,却有相同的主要特性:

1)除了乘法因子 k 的量纲不同外,两种信息和两种熵的数学表达式相同,都由系统的概率分布函数的泛函表述之,都可表示系统的有序度和无序度.

2)两种信息和两种熵各自都遵守数学形式类似的演化方程,即方程(15)(17)(18)和(22)(23)(24)、方程(8)(10)(11)和(1)(2)(3),都表明统计信息密度的时间变化率是由其漂移、扩散和减损三者共同引起的,统计信息熵密度的时间变化率是由其漂移、扩散和产生三者共同引起的.

3)两种信息减损率公式的形式类似,两种熵产生率公式形式亦类似.

两种信息和两种熵的各自差异如下:

1)物理信息和物理熵讨论极大数量的粒子态,其基础具有微观性,适用于统计热力学系统.Shannon 信息和信息熵通常研究较少数的状态,其态变量具有介观性和宏观性,可应用于包括经济学和社会学的多种自然科学和社会科学系统.

2)Shannon 信息和信息熵的演化方程的扩散项和产生项中的梯度,既可是对位置亦可是对速度和其他任何态变量的.物理信息和物理熵的演化方程的扩散项和产生项中的梯度,则仅是对位置而不是对速度和其他任何态变量的.两种信息减损率公式中和两种熵产生率公式中各自的梯度的差异亦如此.

由此可见,Shannon 信息和物理信息、信息熵和物理熵,两种信息和两种熵中各自的前者都可看成是其后者的推广,而其差异则可互为补充.正因此,本文将两种信息和两种熵各自统一称之为统计信息和统计熵.至于它们各自能否完全统一,有待进一步论证.

8.结 论

统计信息(熵),包含 Shannon 信息(熵)和统计物理信息(熵)两种.非平衡统计信息理论,其核心内容是探求描述信息演化规律的信息(熵)演化方程.本文推导出了两种信息(熵)演化方程,数学形式类似.它们都指明信息(熵)密度随时间的变化率是由其漂移、扩散和减损(产生)三者引起的.进而给出了统计信息减损率、统计熵产生率、信息(熵)流、信息(熵)扩散、动态互信息和动态信道容量的表达式,证明了非平衡系统的总的及内部的统计信息与统计熵之和的时间变化率为零.当信息(熵)密度不随时间变化且无信息(熵)流和信息(熵)减损(产生)时,非平衡态统计信息理论就还原为平衡态统计信息理论.所有这些理论结果都是从信息(熵)演化方程出发统一推导出的,未增补新的假设.

[1] Shannon C E 1948 *Bell Sys. Tech. J.* **27** 379,623

[2] Cover T M and Thomas J A 1991 *Elements of Information Theory* (New York:John Wiley & Sons)

[3] Chang J 1993 *Elements of Information Theory* (Beijing: Tsinghua University Press) (in Chinese) [常迥 1993 信息理论基础(北京:清华大学出版社)]

[4] Jaynes E T 1957 *Phys. Rev.* **106** 620

[5] Kapur J N and Kesavan H K 1988 *Entropy Optimization Principle with Application* (San Diego:The MIT Press)

[6] Brillouin L 1962 *Science and Information Theory* (New York: Academic Press)

[7] Weber B H Depew D J and Smith J D Ed. 1988 *Entropy, Information and Evolution* (Cambridge:The MIT Press)

[8] Zurek W H Ed. 1990 *Complexity, Entropy and the Physics of Information* (Redwood City, California :Addison-Wesley)

[9] Haken H 1988 *Information and Self-Organization* (Berlin:Springer-Verlag)

[10] Ingarden H S Kossakowski A and Ohya M 1997 *Information Dynamics and Open System* (Dordrecht:Kluwer Academic Publishers)

[11] Atmanspacher H and Scheingraber H Ed 1991 *Information Dynamics* (New York:Plenum Press)

[12] Barwise J 1997 *Information Flow* (Cambridge:Cambridge University Press)

[13] Weenig W H 1991 *Information Diffusion and Persuation in Communication Networks* (Leiden:Leiden University)

[14] Xing X S 1996 *Science in China* Ser A **26** 617

[15] Xing X S 1998 *Int. J. Mod. Phys.* B **12** 2005

[16] Xing X S 2000 *Chinese Science Bulletin* **45** 1235

[17] Xing X S 2003 *Acta Phys. Sin.* **52** 2969 (in Chinese) [邢修三 2003 物理学报 **52** 2969]

[18] Wu D J, Cai L and Ke S Z 1994 *Phys. Rev.* E **50** 2496

[19] Huang C F 1997 *Fuzzy Sets and System* **91** 69

[20] Haken H 1983 *Synergetics* (Berlin :Springer-Verlag)

Nonequilibrium statistical information theory

Xing Xiu-San

(*Department of Physics*, *Beijing Institute of Technology*, *Beijing* 100081, *China*)

(Received 30 June 2003; revised manuscript received 20 August 2003)

Abstract

In this paper, we propose a nonequilibrium statistical information theory, whose kernel is information (entropy) evolution equation describing information evolution law. A nonlinear evolution equation of Shannon information (entropy) is derived. The statistical physical information is introduced and its nonlinear evolution equation is derived. Both of these two information (entropy) evolution equations show that the temporal change rate of statistical information (entropy) density originates together from their drift, diffusion and dissipation (production) in coodinate space (and state variable space). The expressions of drift information flow and diffusion information flow, the concise formulas of statistical entropy production rate and statistical information dissipation are given. The statistical information dissipation (or increase) rate being equal to its statistical entropy production (or decrease) rate of the dynamic system, the information diffusion and information dissipation occuring at the same time are proved. The dynamic mutual information and dynamic channel capacity reflecting the dynamic dissipative character in transmission process is presented. The similarities and dissimilarities between Shannon information (entropy), its evolution equation and physical information (entropy), its evolution equation are discussed.

Keywords: statistical information (entropy) evolution equation, statistical entropy production rate, statistical information dissipation rate, information (entropy) flow, information (entropy) diffusion, dynamic mutual information, dynamic channel capacity

PACC: 0520, 0540, 8000