

Aktivierung von Neuronenpopulationen und Diffusionsprozesse

Inhaltsverzeichnis

1	Diffusionen	1
1.1	Random walk	1
1.2	Ficksche Diffusion	2
1.3	Markov-Prozesse	3
1.4	Taylor-Entwicklung in 2 Dimensionen	4
1.5	Die Differentielle Chapman-Kolmogoroff Gleichung	4
1.6	Die Fokker-Planck-Gleichung für Diffusionsprozesse	6
1.7	Spezialfall I: der Wiener-Prozess	8
1.8	Spezialfall II: der Ornstein-Uhlenbeck-Prozess	9
1.9	Stochastische Differentialgleichungen	10
1.9.1	Die Langevin-Gleichung	11
1.9.2	Definition des stochastischen Integrals	13
1.9.3	Stochastische Differentialgleichungen	14
1.10	Beispiele	15
1.10.1	Drift und Diffusion unabhängig von x	15
1.10.2	Das Integrate-and-fire-Modell.	16
2	Populationsaktivität	18

1 Diffusionen

1.1 Random walk

Zu festen Zeiten Δt bewegt sich ein Teilchen um einen festen Betrag Δx vorwärts oder rückwärts, jeweils mit der Wahrscheinlichkeit $p = 1/2$. Nach n Schritten befindet sich das Teilchen zwischen $-n\Delta x$ und $n\Delta x$, wenn der Startpunkt in $x = 0$ gelegt wird.

Gesucht ist die Wahrscheinlichkeit $p(m, n)$ dafür, dass der Ort des Teilchens nach n Schritten, also zur Zeit $n\Delta t$, gleich $m\Delta x$ ist. Dazu werde angenommen, dass es insgesamt a Schritte nach rechts und b Schritte nach links getan hat. Dann muß

$$m = a - b, \quad a + b = n, \quad a = \frac{n + m}{2}, \quad b = n - a \tag{1}$$

gelten. Die Anzahl möglicher Pfade bis $m\Delta x$ ist

$$\frac{n!}{a!b!} = \frac{n!}{a!(n-a)!} \tag{2}$$

Die Anzahl möglicher Pfade mit insgesamt n Schritten ist gleich 2^n , so dass die Wahrscheinlichkeit $p(m, n)$ durch

$$p(m, n) = \frac{1}{2^n} \frac{n!}{a!(n-a)!} \quad (3)$$

ist. $p(m, n)$ ist gerade die Binomialverteilung mit dem Parameter $1/2$.

Es werde nun der Fall betrachtet, dass n groß ist, so dass $n \pm m$ ebenfalls groß ist. Dann gilt asymptotisch

$$n! \approx (2\pi n)^{1/2} n^n e^{-n}, \quad n \rightarrow \infty. \quad (4)$$

Dann läßt sich zeigen, dass die Approximation

$$p(m, n) \approx \sqrt{\frac{2}{n\pi}} \exp\left[-\frac{m^2}{2n}\right], \quad m \gg 1, \quad n \gg 1 \quad (5)$$

gilt; diese Approximation ist schon für $n > 6$ bemerkenswert genau.

Es sei nun

$$x = m\Delta x, \quad t = n\Delta t. \quad (6)$$

Weiter gelte $\Delta t \rightarrow 0$, $\Delta x \rightarrow 0$, $m, n \rightarrow \infty$ für x und t endlich. Die Wahrscheinlichkeit $p(m, n)$ geht dann gegen Null, da die Anzahl der Punkte mit $\Delta x \rightarrow 0$ gegen unendlich strebt. Es werde deshalb

$$u =_{def} \frac{p}{2\Delta x} \quad (7)$$

gesetzt, wobei p die Wahrscheinlichkeit ist, dass das Teilchen zur Zeit t im Intervall $(x, x + \Delta x)$ ist. Aus (6) folgt $m = x/\Delta x$ und $n = t/\Delta t$. Dann hat man von (5)

$$\frac{p(x/\Delta x, t/\Delta t)}{2\Delta x} \approx \sqrt{\frac{\Delta t}{2\pi t(\Delta x)^2}} \exp\left[-\frac{x^2}{2t(\Delta x)^2}\right]. \quad (8)$$

Man macht nun die

Annahme:

$$\lim_{\Delta x, \Delta t \rightarrow 0} \frac{(\Delta x)^2}{2\Delta t} \rightarrow D \neq 0. \quad (9)$$

Dann erhält man

$$u(x, t) = \lim_{\Delta x, \Delta t \rightarrow 0} \frac{p(x/\Delta x, \Delta t/t)}{2\Delta x} = \sqrt{\frac{1}{4\pi Dt}} \exp\left[-\frac{x^2}{4Dt}\right]. \quad (10)$$

Hierin heißt D der *Diffusionskoeffizient* oder kurz *Diffusivität* der Teilchen; D hat die Dimension [Länge²/Zeit]. D mißt die Effizienz, mit der Teilchen von einer hohen zu einer niedrigen Dichte übergehen. Im Blut haben die Hämoglobinmoleküle eine Diffusivität von $10^{-7} \text{cm}^2 \text{sek}^{-1}$, während sie für Sauerstoff den Wert $10^{-5} \text{cm}^2 \text{sek}^{-1}$ ist.

1.2 Ficksche Diffusion

Es sei J der Fluß des Materials (der Teilchen: Zellen, Moleküle, etc). Es wird angenommen, dass J proportional zum Gradienten der Konzentration des Materials ist. In einer Dimension hat man dann

$$J \propto -\frac{\partial c}{\partial x}, \quad J = -D \frac{\partial c}{\partial x}. \quad (11)$$

Hierin ist $c(x, t)$ die Konzentration am Ort x zur Zeit t . Das Minuszeichen bedeutet, dass das Material von einer hohen zu einer niedrigen Konzentration transportiert wird.

Annahme: Die Rate der Veränderung des Betrags an Material in einem Bereich ist gleich der Rate des Flusses (*flux*) durch die Grenze des Bereichs (x_0, x_1) , plus der Menge, die innerhalb des Bereichs erzeugt wird.

Es werde kein Material innerhalb des Bereichs erzeugt. Dann gilt

$$\frac{\partial}{\partial t} \int_{x_0}^{x_1} c(x, t) dx = J(x_0) - J(x_1). \quad (12)$$

$\partial c / \partial t$ ist die Rate der Veränderung von c zur Zeit t , das Integral ist die ?

Setzt man $x_1 = x_0 + \Delta x$ und Δx hinreichend klein erhält man insbesondere

$$\frac{\partial}{\partial t} \int_{x_0}^{x_0 + \Delta x} c(x, t) dx \approx \Delta x \frac{\partial c(x_0, t)}{\partial t} = J(x_0) - J(x_0 + \Delta x), \quad (13)$$

so dass

$$\frac{\partial c}{\partial t} = \frac{J(x_0) - J(x_0 + \Delta x)}{\Delta x},$$

und für $\Delta x \rightarrow 0$ erhält man wegen (11)

$$\frac{\partial c}{\partial t} = -\frac{\partial J}{\partial x} = -D \frac{\partial^2 c}{\partial x^2}, \quad (14)$$

dh die Veränderung von $c(x, t)$ zur Zeit t ist proportional zur Veränderung des Gradienten ($\partial c / \partial x$ ist der Gradient, die zweite Ableitung ist die Veränderung des Gradienten).

Setzt man insbesondere

$$c(x, 0) = Q\delta(x), \quad (15)$$

so erhält man als Lösung für (14)

$$c(x, t) = \frac{Q}{2\sqrt{\pi Dt}} e^{-x^2/(4Dt)}, \quad t > 0. \quad (16)$$

Für $Q = 1$ wird diese Gleichung identisch mit der Gleichung (10) für den Random Walk.

1.3 Markov-Prozesse

Eine zufällige Veränderliche X nehme zum Zeitpunkt t den Wert x an; allgemein sei $(x_1, t_1), \dots, (x_n, t_n)$ eine Folge von Zeit-Zustands-Paaren; zum Zeitpunkt t_j nimmt X den Wert x_j an. Die Folge ist die Realisierung eines stochastischen Prozesses. Es wird die bedingte Wahrscheinlichkeit für (x_n, t_n) betrachtet, gegeben die Paare $(x_1, t_1), \dots, (x_{n-1}, t_{n-1})$. Gilt

$$P_n(x_n, t_n | x_{n-1}, t_{n-1}; \dots; x_1, t_1) = P_2(x_n, t_n | x_{n-1}, t_{n-1}), \quad (17)$$

so ist der Prozess ein Markov-Prozess. Die Vergangenheit, dh die Paare $(x_{n-2}, t_{n-2}; \dots; x_1, t_1)$, kann also vernachlässigt werden; in diesem Sinne ist der Prozess gedächtnislos.

Markov-Prozesse sind eindeutig durch $p_1(x, t)$ und $p_2(x_2, t_2 | x_1, t_1)$ bestimmt. So gilt für $t_3 > t_2 > t_1$

$$\begin{aligned} p_3(x_3, t_3 | x_2, t_2; x_1, t_1) &= p_3(x_3, t_3 | x_2, t_2; x_1, t_1) p_2(x_2, t_2 | x_1, t_1) \\ &= p_2(x_3, t_3 | x_2, t_2) p_2(x_2, t_2 | x_1, t_1) p_1(x_1, t_1). \end{aligned} \quad (18)$$

Chapman-Kolmogoroff-Gleichung: Der Prozess starte ind (x_1, t_1) . Um in t_3 nach x_3 zu gelangen, kann der Prozess über alle möglichen x -Werte laufen; um also die Wahrscheinlichkeit $p_2(x_3, t_3|x_1, t_1)$ zu bestimmen, muß (18) über alle x_2 integriert werden. Man erhält

$$p(x_1, t_1|x_3, t_3) = \int p(x_1, t_1|x_2, t_2)p(x_2, t_2|x_3, t_3)dx_2, \quad t_1 \geq t_2 \geq t_3 \quad (19)$$

dies ist die *Chapman-Kolmogoroff-Gleichung*.

Kontinuierlicher Markov-Prozess: Mit Wahrscheinlichkeit 1 sind die Trajektorien eines Markov-Prozesses stetig, wenn

$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{|x-z|<\varepsilon} p(x, t + \Delta t|z, t)dx = 0 \quad (20)$$

gilt. Man spricht dann von einem kontinuierlichen Markov-Prozess.

1.4 Taylor-Entwicklung in 2 Dimensionen

$$\begin{aligned} f(x+h, y+k) &= f(x, y) + [hf_x(x, y) + kf_y(x, y)] + \\ &+ \frac{1}{2!}[h^2f_{xx}(x, y) + k^2f_{yy}(x, y) + 2hkf_{xy}(x, y)] + \dots \end{aligned}$$

Man kann diese Formel vektoriell schreiben; das erleichtert den Übergang zu $n > 2$ Variablen. Setzt man $x_1 = x$, $x_2 = y$, und setzt man $\vec{x} = (x_1, x_2)'$, $\vec{z} = (z_1, z_2)' = (x_1 + h, x_2 + k)'$, so ergibt sich die allgemeine Formel

$$\begin{aligned} f(\vec{x}) &= f(\vec{z}) + \sum_i \frac{\partial f(\vec{z})}{\partial z_i} (x_i - z_i) \\ &+ \sum_{i,j} \frac{1}{2} \frac{\partial^2 f(\vec{z})}{\partial z_i \partial z_j} (x_i - z_i)(x_j - z_j) + |\vec{x} - \vec{z}|^2 R(\vec{x}, \vec{z}) \end{aligned} \quad (21)$$

wobei $|R(\vec{x}, \vec{z})| \rightarrow 0$ mit $|\vec{x} - \vec{z}| \rightarrow 0$. Natürlich kann \vec{x} einen beliebigen n -dimensionalen Vektor, $n > 1$ bezeichnen (also kann $n > 2$ sein).

1.5 Die Differentielle Chapman-Kolmogoroff Gleichung

Gegeben sei eine Funktion $f(\vec{x})$, und es sei $p(\vec{x}, t|\vec{y}, t')$. Dann ist

$$E(f) = \int f(\vec{x})p(\vec{x}, t|\vec{y}, t')d\vec{x}$$

der Erwartungswert von $f(\vec{x})$. Mit ∂_t werde die partielle Ableitung nach t bezeichnet. Die partielle Ableitung

$$\partial_t E(f) = \partial_t \int f(\vec{x})p(\vec{x}, t|\vec{y}, t')d\vec{x} \quad (22)$$

bezeichnet dann die die differentielle Veränderung des Erwartungswertes mit der Zeit. Aus der Definition des Differentialquotienten ergibt sich dann

$$\partial_t E(f) = \lim_{\Delta t \rightarrow 0} \left[\int f(\vec{x}) \frac{p(\vec{x}, t + \Delta t|\vec{y}, t') - p(\vec{x}, t|\vec{y}, t')}{\Delta t} d\vec{x} \right], \quad (23)$$

und dieser Ausdruck kann wieder in der Form

$$\partial_t E(f) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \left[\int f(\vec{x}) p(\vec{x}, t + \Delta t | \vec{y}, t') d\vec{x} - \int f(\vec{x}) p(\vec{x}, t | \vec{y}, t') d\vec{x} \right] \quad (24)$$

Auf das erste Integral auf der rechten Seite dieser Gleichung kann nun die Chapman-Kolmogoroff Gleichung (19) angewendet werden, d.h.

$$p(\vec{x}, t + \Delta t | \vec{y}, t') = \int p(\vec{x}, t + \Delta t | \vec{z}, t) p(\vec{z}, t | \vec{y}, t') d\vec{z}, \quad t' \leq t \leq t + \Delta t. \quad (25)$$

Im rechten Integral kann man \vec{x} durch \vec{z} ersetzen; es wird ja über *alle* \vec{x} integriert. Dann erhält man

$$\partial_t E(f) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \left[\iint f(\vec{x}) p(\vec{x}, t + \Delta t | \vec{z}, t) p(\vec{z}, t | \vec{y}, t') d\vec{z} d\vec{x} - \int f(\vec{z}) p(\vec{z}, t | \vec{y}, t') d\vec{z} \right]. \quad (26)$$

Es werden nun drei Bedingungen eingeführt, die für alle $\varepsilon > 0$ gelten sollen:

- (i) $\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} p(\vec{x}, t + \Delta t | \vec{z}, t) = W(\vec{x} | \vec{z}, t)$, gleichförmig in \vec{x}, \vec{z} und t für $|\vec{x} - \vec{z}| \geq \varepsilon$,
- (ii) $\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{|\vec{x} - \vec{z}| < \varepsilon} (x_i - z_i) p(\vec{x}, t + \Delta t | \vec{z}, t) d\vec{x} = A_i(\vec{z}, t) + O(\varepsilon)$;
- (iii) $\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{|\vec{x} - \vec{z}| < \varepsilon} (x_i - y_i)(x_j - y_j) p(\vec{x}, t + \Delta t | \vec{z}, t) d\vec{x} = B_{ij}(\vec{z}, t) + O(\varepsilon)$

Erläuterungen:

1. Der Term $O(\varepsilon)$ steht für "in der Ordnung von". Sind ζ_n und z_n zwei Folgen und gilt $|\zeta_n/z_n| < K$, wenn $n > n_0$ und K ist eine Konstante unabhängig von n , dann ist $\zeta_n = o(z_n)$. So ist $(15n + 19)/(1 + n^3) = O(1/n^2)$, denn

$$\left[\frac{15n + 19}{1 + n^3} \right] \frac{n^2}{1} = \frac{15 + 19/n}{1/n + n^2} \rightarrow \frac{15}{n^2} \rightarrow 0,$$

dhvon einem bestimmten n ab ist der Quotient kleiner als jede vorgegebene Schranke K . Gilt insbesondere $\lim \zeta_n/z_n = 0$, so ist $\zeta_n = o(z_n)$, was für das Beispiel ebenfalls gilt. $O(\varepsilon)$ steht also für einen Rest R , für den R/ε kleiner als eine vorgeschriebene Schranke ist. Da ε selbst beliebig klein ist, muß also R noch viel kleiner sein, dh $O(\varepsilon)$ steht in diesem Sinne für eine vernachlässigbare Größe, wenn nur ε hinreichend klein ist.

2. $A_i(\vec{z}, t)$ steht offenbar für den Erwartungswert der Differenz $x_i - y_i$ der i -ten Komponenten von \vec{x} und \vec{y} im Bereich $|\vec{x} - \vec{y}| < \varepsilon$, und zwar für $\Delta t \rightarrow 0$, wenn sich der Wert von t nur infinitesimal ändert. $A_i(\vec{z}, t)$ bezeichnet die *Drift* des Prozesses.
3. B_{ij} für die entsprechende Kovarianz der Differenzen $x_i - y_i$ und $x_j - y_j$ bei infinitesimaler Veränderung von t . B_{ij} kennzeichnet die *Diffusion* des Prozesses.

Man kann nun die allgemeine Taylor-Entwicklung von $f(\vec{z})$ in die Gleichung (26) einsetzen. In dieser Gleichung wird über alle \vec{x} und \vec{y} integriert. Man kann den Bereich von \vec{x} (und damit von \vec{y}) in die Bereiche $|\vec{x} - \vec{y}| < \varepsilon$ und $|\vec{x} - \vec{y}| \geq \varepsilon$ aufteilen. Berücksichtigt man dabei die Bedingungen (i) - (iii), gelangt man schließlich zur

Differentiellen Chapman-Kolmogoroff-Gleichung:

$$\begin{aligned} \partial_t p(\vec{z}, t | \vec{y}, t') &= - \sum_i \frac{\partial}{\partial z_i} [A_i(\vec{z}, t) p(\vec{z}, t | \vec{y}, t')] \\ &+ \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial z_i \partial z_j} [B_{ij}(\vec{z}, t) p(\vec{z}, t | \vec{y}, t')] \\ &+ \int [W(\vec{z} | \vec{x}, t) p(\vec{x}, t | \vec{y}, t) - W(\vec{x} | \vec{z}, t) p(\vec{z}, t | \vec{y}, t')] d\vec{x} \end{aligned} \quad (27)$$

Die Ausdrücke W beziehen sich dabei auf den Bereich $|\vec{x} - \vec{y}| \geq \varepsilon$, also auf nicht-infinitesimale Veränderungen (Sprünge) des Zustandsvektors. Lösungen für diese Gleichung existieren, wenn die *Initialbedingung*

$$p(\vec{z}, t|\vec{y}, t) = \delta(\vec{y} - \vec{z}), \quad (28)$$

δ die Dirac-Delta-Funktion, erfüllt ist.

1.6 Die Fokker-Planck-Gleichung für Diffusionsprozesse

Diese Gleichung ergibt sich aus der differentiellen Chapman-Kolmogoroff-Gleichung, wenn das Integral in (27) gleich Null ist. Der Prozess macht dann keine Sprünge, dh die Trajektorien sind stetig. Man schreibt

$$\frac{\partial p(\vec{z}, t|\vec{y}, t')}{\partial t} = - \sum_i \frac{\partial}{\partial z_i} [A_i(\vec{z}, t)p(\vec{z}, t|\vec{y}, t')] + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial z_i \partial z_j} [B_{ij}(\vec{z}, t)p(\vec{z}, t|\vec{y}, t')] \quad (29)$$

Man kann die A_i zu einem Vektor $\vec{A}(\vec{z}, t)$ und die $B_{ij}(\vec{z}, t)$ zu einer Matrix $B(\vec{z}, t)$ zusammenfassen; \vec{A} ist der *Driftvektor* und B ist die *Diffusionsmatrix*. B ist symmetrisch und positiv-semidefinit. Wird ein Prozess durch die Fokker-Planck-Gleichung beschrieben, so existieren für den Zustandsvektor \vec{x} keine Sprünge, so dass die Trajektorien des Prozesses stetige Funktionen der Zeit sind.

Man vergleiche die Fokker-Planck-Gleichung (FPG) (29) mit dem Ausdruck für die Ficksche Diffusion (14), $\partial c/\partial t = -D\partial^2 c/\partial x^2$. Ist $\vec{A} \equiv 0$, so entspricht die FPG der Fickschen Diffusion mit $-D = B$. Setzt man hier wieder $D = (\Delta x)^2/2\Delta t$ mit $\Delta x \rightarrow 0$, $\Delta t \rightarrow 0$, so sieht man, dass $(\Delta x)^2/2\Delta t$ einer Varianz pro Zeiteinheit entspricht. Vergleichbare Größen stehen in der Diagonalen von B , und in den nicht-diagonalen Zellen stehen entsprechende Kovarianzgrößen.

Anmerkungen zum Drift- und Diffusionskoeffizienten: Der Einfachheit halber wird der 1-dimensionale Fall betrachtet. Es ist $p(x, t|x_0, t_0)$ die Wahrscheinlichkeit, dass der Prozess zum Zeitpunkt t in x ist unter der Bedingung, dass er zum Zeitpunkt t_0 in x_0 ist. Man kann für x den Erwartungswert bestimmen:

$$a(x_0, t, t_0) = \int xp(x, t|x_0, t_0)dx. \quad (30)$$

$a(x_0, t, t_0)$ ist die bedingte Erwartung für x , dh der Erwartungswert für x zur Zeit t , wenn in t_0 der Zustand x_0 war. Dementsprechend kann man die bedingte Varianz

$$b(x_0, t, t_0) = \int (x - a)^2 p(x, t|x_0, t_0) dx \quad (31)$$

definieren. Es sei $t = t_0 + \Delta t$. Für hinreichend kleine Δt gelten dann die Reihenentwicklungen

$$a(x_0, t_0 + \Delta t, t_0) \approx x_0 + \Delta t \frac{\partial}{\partial t} a(x_0, t, t_0)|_{t=t_0} = x_0 + \Delta t \eta(x_0, t_0) \quad (32)$$

$$b(x_0, t_0 + \Delta t, t_0) \approx \Delta t \frac{\partial}{\partial t} b(x_0, t, t_0)|_{t=t_0} = \Delta t \sigma^2(x_0, t_0) \quad (33)$$

(32) entspricht der Taylor-Entwicklung $f(t + \Delta t) = f(t) + \Delta t f'(t) + \dots$ für eine beliebige Funktion f , falls diese in eine Reihe entwickelbar ist, dh falls f' , f'' etc existieren. $f(t)$ entspricht $a(x_0, t_0 + \Delta t, t_0)$, und für $\Delta t = 0$ erhält man $p(x, t_0|x_0, t_0)$. Da der Prozess stetig ist, macht er keine Sprünge, so dass $x = x_0$ sein muß; in (31) ist also $p(x, t|x_0, t_0) = \delta(x - x_0)$, so dass

$a = \int x\delta(x-x_0)dx = x_0$ ist. Für $\Delta t = 0$ ergibt sich für (31) $b(x_0, t_0, t_0) = \int (x-a)\delta(x-x_0)dx = x_0 - a$. Aber $a = x_0$, so dass $b(x_0, t_0, t_0) = 0$. Man bemerke, dass nach (32) und (32)

$$\eta(x_0, t_0) = \frac{\partial}{\partial t} a(x_0, t, t_0)|_{t=t_0} \quad (34)$$

$$\sigma^2(x_0, t_0) = \frac{\partial}{\partial t} b(x_0, t, t_0)|_{t=t_0}, \quad (35)$$

dh $\eta(x_0, t_0)$ ist die infinitesimale Veränderung der bedingten Erwartung $a(x, t, t_0)$ mit t , und $\sigma^2(x_0, t_0)$ ist die infinitesimale Veränderung der bedingten Varianz $b(x_0, t_0)$. (32) und (33) korrespondieren zu den Bedingungen (ii) und (iii) für die Gleichung (26).

Die bedingte Dichte $p(x, t; x_0, t_0)$ für $t > t_0$ genügt dann der Fokker-Planck-Gleichung, die jetzt in der Form

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x} [\eta(x, t)p(x, t)] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [\sigma^2(x, t)p(x, t)] \quad (36)$$

geschrieben werden kann (Papoulis (1965), p. 538).

Heuristische Beschreibung des Prozesses: Es gelte wieder die Initialbedingung (28); es soll $p(\vec{z}, t + \Delta t | \vec{y}, t)$ berechnet werden. Ist nun Δt klein, so wird $p(\vec{z}, t + \Delta t | \vec{y}, t)$ immer noch einen scharfen "Peak" um \vec{y} bilden, so dass die Ableitungen von A_i und B_{ij} vernachlässigbar sind im Vergleich zu der von p . Vernachlässigt man noch die zeitliche Abhängigkeit von A_i und B_{ij} für kleine $t - t'$, so erhält man

$$\frac{\partial p(\vec{z}, t | \vec{y}, t')}{\partial t} = -\sum_i A_i(\vec{z}, t) \frac{\partial p(\vec{z}, t | \vec{y}, t')}{\partial z_i} + \frac{1}{2} \sum_{i,j} B_{ij}(\vec{y}, t) \frac{\partial^2 p(\vec{z}, t | \vec{y}, t')}{\partial z_i \partial z_j} \quad (37)$$

Diese Gleichung kann gelöst werden; man erhält

$$p(\vec{z}, t + \Delta t | \vec{y}, t) = C \exp \left[-\frac{1}{2} \frac{[\vec{z} - \vec{y} - \vec{A}(\vec{y}, t)\Delta t]' B^{-1}(\vec{y}, t) [\vec{z} - \vec{y} - \vec{A}(\vec{y}, t)\Delta t]}{\Delta t} \right] \quad (38)$$

mit

$$C = \frac{1}{(2\pi)^{n/2}} \sqrt{\frac{\det B(\vec{y}, t)}{\Delta t}}, \quad (39)$$

wobei $\det B$ die Determinante von B ist. (38) ist die Gleichung der Dichte eines Gauss-verteiltern Vektors $\vec{z} - \vec{y}$, mit $\vec{z} = \vec{z}(t) = \vec{y}(t + \Delta t)$. Der Erwartungswert dieses Vektors ist $\vec{y} + \vec{A}(\vec{y}, t)\Delta t$, und die Varianz-Kovarianz-Matrix der Komponenten ist $B(\vec{y}, t)$. Nun ist

$$\frac{1}{\Delta t} B^{-1}(\vec{y}, t) / \Delta t = (\Delta t B(\vec{y}, t))^{-1},$$

dh die Komponenten von $\vec{y}(t + \Delta t) - \vec{y}(t)$ haben die Varianz-kovarianz-Matrix $B(\vec{y}, t)\Delta t$. Dementsprechend kann man

$$\vec{y}(t + \Delta t) = \vec{y}(t) + \vec{A}[\vec{y}(t), t]\Delta t + \vec{\eta}(t)\sqrt{\Delta t} \quad (40)$$

schreiben. Hier ist $\vec{\eta}(t)$ ein Gauss-verteilter zufälliger Vektor mit Erwartungswert und Varianz-Kovarianz

$$\langle \vec{\eta}(t) \rangle = 0, \quad \langle \vec{\eta}\vec{\eta}' \rangle = B(\vec{y}, t). \quad (41)$$

Aus (40) läßt sich nun ein Ausdruck für den Differenzenquotienten von \vec{y} herleiten:

$$\frac{\vec{y}(t + \Delta t) - \vec{y}(t)}{\Delta t} = \vec{A}(\vec{y}, t) + \frac{\vec{\eta}(t)}{\sqrt{\Delta t}}. \quad (42)$$

Die Rate der Veränderung von $\vec{y}(t)$ ist also gleich dem Drift-Term $\vec{A}(\vec{y}, t)$, plus einem zufälligen Term $\vec{\eta}(t)/\sqrt{\Delta t}$. Man sieht, dass der Grenzwert $\lim_{\Delta t \rightarrow 0}$ nicht existiert, da $\vec{\eta}(t)/\sqrt{\Delta t} \rightarrow \infty$ für $\Delta t \rightarrow 0$. Da diese Aussage für beliebige t gilt, folgt, dass $\vec{y}(t)$ zwar überall stetig, aber nirgends differenzierbar ist.

1.7 Spezialfall I: der Wiener-Prozess

Der Prozess ist nach Norbert Wiener benannt, der diesen Prozess zuerst herleitete.

Es sei $n = 1$, so dass $\vec{x} = x = w$, wobei w geschrieben wird, um den Wiener-Prozess anzuzeigen. Die Trajektorien sind zufällige Funktionen $W(t)$. Die Fokker-Planck-Gleichung ist

$$\frac{\partial p(w, t|w_0, t_0)}{\partial t} = \frac{1}{2} \frac{\partial^2}{\partial w^2} p(w, t|w_0, t_0). \quad (43)$$

Es gelte die Anfangsbedingung

$$p(w, t_0|w_0, t_0) = \delta(w - w_0). \quad (44)$$

Der FPG (43) entspricht dann die Funktion

$$p(w, t|w_0, t_0) = \frac{1}{\sqrt{2\pi(t-t_0)}} \exp\left[-\frac{(w-w_0)^2}{2(t-t_0)}\right], \quad (45)$$

wie man durch Differenzieren überprüft. Demnach ist w Gauss-verteilt mit dem Erwartungswert und der Varianz

$$\langle W(t) \rangle = w_0, \quad \langle [W(t) - w_0]^2 \rangle = t - t_0. \quad (46)$$

Der 1-dimensionale Wiener-Prozess wird auch *Brownsche Bewegung* genannt, weil er der von Einstein hergeleitete Lösung für die Brownsche Bewegung entspricht.

Der multivariate Wiener-Prozess ist durch den Vektor $\vec{W}(t) = (W_1(t), \dots, W_n(t))'$ definiert. Die FPG ist jetzt

$$\frac{\partial}{\partial t} p(\vec{w}, t|\vec{w}_0, t_0) = \frac{1}{2} \sum_i \frac{\partial^2}{\partial w_i^2} p(\vec{w}, t|\vec{w}_0, t_0) \quad (47)$$

und hat die Lösung

$$p(\vec{w}, t|\vec{w}_0, t_0) = \frac{1}{[2\pi(t-t_0)]^{n/2}} \exp\left[-\frac{(\vec{w} - \vec{w}_0)^2}{2(t-t_0)}\right] \quad (48)$$

$$= \frac{1}{[2\pi(t-t_0)]^{n/2}} \exp[-(\vec{w} - \vec{w}_0)' \Sigma^{-1} (\vec{w} - \vec{w}_0)] \quad (49)$$

mit $\Sigma^{-1} = \text{diag}[1/2(t-t_0), \dots, 1/2(t-t_0)]$. (49) stellt nur die Beziehung zur üblichen Schreibweise der multivariaten Normalverteilung her. Man hat

$$\langle W(t) \rangle = w_0, \quad \langle [W_i(t) - w_{0i}][W_j(t) - w_{0j}] \rangle = (t - t_0) \delta_{ij} \quad (50)$$

Die Trajektorien des Wiener-Prozesses sind extrem irregulär und sind, wie alle Trajektorien von Prozessen, die durch die FPG beschrieben werden, nicht differenzierbar. Man betrachte dazu noch einmal die Wahrscheinlichkeit

$$P\left[\frac{W(t+h) - W(t)}{h} > k\right] = \frac{1}{\sqrt{2\pi h}} \int_k^\infty \exp(-w^2/h) = 1 \text{ für } h \rightarrow 0. \quad (51)$$

Dies bedeutet, dass der Differentialquotient an jedem Punkt fast sicher unendlich ist.

Unabhängigkeit der Inkremente: Der Wiener-Prozess ist der Grundbaustein für die Analyse von Diffusionsprozessen. Deshalb ist die folgende Eigenschaft von besonderem Interesse.

Man betrachtet die Inkremente $\Delta W_i \equiv W(t_i) - W(t_{i-1})$, $\Delta t_i = t_i - t_{i-1}$. Da der Wiener-Prozess ein Markov-Prozess ist kann man

$$p(w_n, t_n; w_{n-1}, t_{n-1}; \dots; w_0, t_0) = \prod_{i=0}^{n-1} p(w_{i+1}, t_{i+1} | w_i, t_i) p(w_0, t_0) \quad (52)$$

Wendet man hierauf (48) an, so erhält man

$$p(w_n, t_n; \dots; w_0, t_0) = \prod_{i=0}^{n-1} \frac{1}{\sqrt{2\pi(t_{i+1} - t_i)}} \exp. \quad (53)$$

Dem entspricht die Dichte

$$p(\Delta w_n; \dots; \Delta w_1; w_0) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\Delta t_i}} \exp \left[-\frac{\Delta w_i^2}{2\Delta t_i} \right] p(w_0, t_0). \quad (54)$$

Dieser Ausdruck ist aber gleichbedeutend damit, dass die Inkremente Δw_i alle stochastisch unabhängig sind. Diese Eigenschaft wird bei der Definition des stochastischen Integrals ausgenutzt.

Autokorrelation: Die Autokorrelationsfunktion ist durch

$$\langle W(t)W(s) | [w_0, t_0] \rangle = \int w_1 w_2 p(w_1, t; w_2, s | w_0, t_0) dw_1 dw_2 \quad (55)$$

definiert. Dies ist das mittlere Produkt von $W(t)$ und $W(s)$ unter der Bedingung, dass der Prozess in (w_0, t_0) startet. Nimmt man nun $t > s$ an, so hat man

$$\langle W(t)W(s) | [w_0, t_0] \rangle = \langle [W(t) - W(s)]W(s) \rangle + \langle [W(s)]^2 \rangle. \quad (56)$$

Hierbei ist $\langle [W(t) - W(s)]W(s) \rangle = 0$ wegen der Unabhängigkeit der Zuwächse (Inkremente), und für den zweiten Term erhält man wegen (46) $t - t_0$, so dass man allgemein

$$\langle W(s)W(t) | [w_0, t_0] \rangle = \min[t - t_0, t_0 - t] + w_0^2, \quad (57)$$

und diese Formel gilt für $t > s$ und $t < s$.

Der Wiener-Prozess hat einige schrille Eigenschaften und wird eigentlich erst interessant in Zusammenhang mit stochastischen Differentialgleichungen. Ein realistischerer Prozess ist der Ornstein-Uhlenbeck-Prozess.

1.8 Spezialfall II: der Ornstein-Uhlenbeck-Prozess

Es werde die FPG

$$\partial_t p = \partial_x(kxp) + \frac{1}{2}D\partial^2 p \quad (58)$$

betrachtet, mit $p = p(x, t | x_0, 0)$. Hier tritt also ein linearer Drift-Term kx auf. Es werde eine stationäre Lösung gesucht, so dass $\partial_t p = 0$ gilt. Dann hat man

$$0 = \partial_x(kxp) + \frac{1}{2}D\partial^2 p = \partial_x[kxp + \frac{1}{2}D\partial_x p]. \quad (59)$$

Integriert man die rechte Seite einmal in bezug auf x , so erhält man

$$\left[kxp + \frac{1}{2} \partial_x p \right]_{-\infty}^x = 0. \quad (60)$$

Damit p eine Wahrscheinlichkeit ist, muß p für $x \rightarrow \infty$ zusammen mit der Ableitung $\partial_x p$ verschwinden; dann erhält man

$$\frac{1}{p} \partial_x p = -\frac{2kx}{D}, \quad (61)$$

woraus

$$p_s(x) = \frac{1}{\sqrt{\pi D/k}} \exp(-kx^2/D) \quad (62)$$

folgt. x ist also Gauss-verteilt mit dem Mittelwert 0 und der Varianz $D/2k$. Damit ist der Ornstein-Uhlenbeck-Prozess als einstationärer Gauss-Prozess charakterisiert.

Autokorrelation und Korrelationszeit: Die Autokorrelationsfunktion des OU-Prozesses ist

$$R(\tau) = \langle X(t)X(s) \rangle = \frac{D}{2k} \exp(-k|t-s|), \quad \tau = |t-s|. \quad (63)$$

$R(\tau)$ hängt also nicht von den Werten von t und s selbst, sondern nur von der Differenz $t-s$ ab, was für stationäre Prozesse charakteristisch ist.

Es sei $\tau_0 = 1/k$. Dann ist

$$R(|t-s|) = -\frac{D}{2k} \exp\left[-\frac{1}{\tau_0}|t-s|\right],$$

so dass insbesondere

$$R(\tau_0) = \frac{D}{2k} \exp(-1) = \frac{D}{2ke}. \quad (64)$$

Für $\tau > \tau_0$ wird $R(\tau)$ schnell klein, dh vernachlässigbar. Deshalb heißt $\tau_0 = 1/k$ die *Korrelationszeit* (*correlation time*). Allgemein wird die Korrelationszeit definiert durch

$$\tau_0 = \int_0^\infty \frac{\langle X(t), X(0) \rangle_s}{\text{Var}(X)} dt, \quad (65)$$

und dieser Ausdruck ist unabhängig von der speziellen Verteilungsfunktion des Prozesses, gilt also nicht nur für den OU-Prozess.

1.9 Stochastische Differentialgleichungen

Es werde noch einmal die Gleichung (40), also

$$\vec{y}(t + \Delta t) = \vec{y}(t) + \vec{A}[\vec{y}(t), t] \Delta t + \vec{\eta}(t) \sqrt{\Delta t}.$$

betrachtet; sie gibt die Veränderung von $y(t)$ in Abhängigkeit von \vec{A} an. Die Betrachtung von (42), dh Subtraktion von $\vec{y}(t)$, Division durch Δt , dh

$$\frac{1}{\Delta t} (\vec{y}(t + \Delta t) - \vec{y}(t)) = \vec{A}[\vec{y}(t), t] + \vec{\eta}(t) \frac{\sqrt{\Delta t}}{\Delta t} = \vec{A}[\vec{y}(t), t] + \vec{\eta}(t) \frac{1}{\sqrt{\Delta t}}$$

zeigte, dass der Grenzwert $d\vec{y}(t)/dt$ für $\Delta t \rightarrow 0$ nicht existiert. Existierte er, könnte man für $\Delta t \approx dt$ wegen $\vec{y}(t + dt) \approx \vec{y}(t) + dt \vec{y}'(t)$, $\vec{y}'(t) = d\vec{y}(t)/dt$ zu einer Differentialgleichung für die Trajektorien \vec{y} gelangen, die stochastische Differentialgleichungen wären, da ja die zufällige Größe $\vec{\eta}$ darin aufträte. Eine solche Differentialgleichung würde es ermöglichen, den Verlauf solcher Trajektorien zu betrachten. Will man solche Gleichungen aufstellen, muß man also anders vorgehen.

1.9.1 Die Langevin-Gleichung

Gegeben sei die Differentialgleichung

$$\frac{dv(t)}{dt} = a(v, t) + b(v, t)\xi(t), \quad (66)$$

wobei $a(v, t)$ und $b(v, t)$ gegebene Funktionen seien und $\xi(t)$ eine Trajektorie eines schnell fluktuierenden Zufallsprozesses seien. "Schnell fluktuierend" soll dabei bedeuten, dass $\xi(t)$ und $\xi(t')$ für $t \neq t'$ stochastisch unabhängig voneinander sind, dh es soll

$$\langle \xi(t), \xi(t') \rangle = \delta(t - t') \quad (67)$$

gelten. Außerdem soll $\langle \xi(t) \rangle = 0$ sein für alle t ; diese Forderung kann immer aufgestellt werden, da $\mu(t) = \langle \xi(t) \rangle \neq 0$ in die Funktion $a(v, t)$ absorbiert werden kann.

Anmerkung: Die Gleichung (66) wurde zuerst von Langevin¹ im Zusammenhang mit der Frage nach der Brownschen Bewegung aufgestellt². Die Beobachtung, dass sich die Teilchen, in die Pollen in wäßriger Lösung zerfallen, sich in dauerhafter, unregelmäßiger Bewegung befinden, wurde von Brown im Jahre 1827 gemacht. Die Erklärung des Phänomens kam erst 1900 von Exner, demzufolge die Teilchen durch die Moleküle des Lösungsmittels - Wasser - gestoßen würden. Die Richtungsänderung eines Teilchen ist dabei nicht die Folge eines Stoßes, denn ein Stoß findet im Mittel alle 10^{-10} Sekunden statt, die Resultate dieser Stöße sind also im Einzelnen gar nicht wahrnehmbar. Vielmehr summiert sich die thermische Bewegung zu einer stochastischen Kraft auf, die durch einen zeitabhängigen Zufallsvektor \vec{K}_t beschrieben werden kann. \vec{K}_t repräsentiert den Einfluß vieler Moleküle, der in guter Nähe durch weißes Rauschen abgebildet werden kann. Außerdem legt dieser Sachverhalt nahe, dass \vec{K}_t Gass-verteilt ist. Kraft ist Masse \times Beschleunigung, und Beschleunigung ist die Ableitung der Geschwindigkeit, dv/dt . Dementsprechend kann man

$$m \frac{dv(t)}{dt} = -\gamma v(t) + K_t \quad (68)$$

schreiben. Der Term $-\gamma v(t)$ repräsentiert eine Reibungskraft, die proportional zur Geschwindigkeit ist und die der Kraft entgegen wirkt. (68) wurde von Langevin aufgestellt und heißt dementsprechend Langevin-Gleichung. Der Gleichung entsprechend überträgt sich der zufallscharakter der Kraft K_t auf die momentane Geschwindigkeit $v(t)$. Die Lösung der Gleichung ergibt sich eine Trajektorie v , wobei die Verteilung von v zu jedem Zeitpunkt t mit angegeben werden soll.

Die Gleichung (66) kann natürlich in der Form

$$\frac{dX(t)}{dt} = a(X, t) + b(X, t)\xi(t), \quad (69)$$

geschrieben werden, um an die übliche Schreibweise für Differentialgleichungen (Dgln) anzuknüpfen.

Die Annahme (67) ist unrealistisch und kann nur als eine Idealisierung dienen, denn sie impliziert u.a., dass die Varianz des Prozesses unendlich ist. Eine realistischere Annahme wäre

$$\langle X(t), X(t') \rangle = \frac{\gamma}{2} e^{-\gamma|t-t'|}. \quad (70)$$

Für $\gamma \rightarrow 0$ strebt die rechte Seite gegen eine δ -Funktion.

¹ Langevin, Paul (1872-1946)

² Brown, R.: A Brief Account of Microscopical Observations Made in the Months on June, July, and August, 1827, on the Particles Contained in the Pollen of Plants; and on the General Existence of Active Molecules in Organic and Inorganic Bodies. Phil. Mag. 4, 161-173, 1828.

Ein alternativer Ansatz geht davon aus, dass (69) eben eine Dgl ist. Damit sie eine Lösung hat, muß

$$u(t) = \int_0^t \xi(t') dt' \quad (71)$$

existieren. Als Integral muß $u(t)$ eine stetige Funktion von t sein. Weiter kann man

$$u(t') = \int_0^t \xi(s) ds + \int_t^{t'} \xi(s) ds = \lim_{\varepsilon \rightarrow 0} \left[\int_0^{t-\varepsilon} \xi(s) ds \right] + \int_t^{t'} \xi(s) ds \quad (72)$$

schreiben. Für alle $\varepsilon > 0$ sind die $\xi(s)$ im ersten Integral unabhängig von denen im zweiten Integral, und wegen der Stetigkeit sind die $u(t)$ und $u(t') - u(t)$ statistisch unabhängig voneinander, und weiter ist $u(t') - u(t)$ unabhängig von $u(t'')$ für alle $t'' < t$. Also ist $u(t')$ vollständig probabilistisch bestimmt durch die Kenntnis von $u(t)$ und nicht durch weiter zurückliegende Werte. Daraus folgt, dass $u(t)$ ein Markov-Prozess ist.

Da die $u(t)$ stetige Funktionen sind, lassen sie sich durch eine Fokker-Planck-Gleichung beschreiben. Dazu müssen der Drift- und der Diffusionskoeffizient bestimmt werden. Es ist

$$\langle u(t + \Delta t) - u_0 | [u_0, t_0] \rangle = \left\langle \int_t^{t+\Delta t} \xi(s) ds \right\rangle = 0 \quad (73)$$

und

$$\begin{aligned} \langle [u(t + \Delta t) - u_0]^2 | [u_0, t_0] \rangle &= \int_t^{t+\Delta t} ds \int_t^{t+\Delta t} ds' \langle \xi(s), \xi(s') \rangle \\ &= \int_t^{t+\Delta t} ds \int_t^{t+\Delta t} ds' \delta(s - s') = \Delta t. \end{aligned} \quad (74)$$

Damit hat man den Drift- und den Diffusionskoeffizienten:

$$A(u_0, t) = \lim_{\Delta t \rightarrow 0} \frac{\langle u(t + \Delta t) - u_0 | [u_0, t_0] \rangle}{\Delta t} = 0 \quad (75)$$

$$B(u_0, t) = \lim_{\Delta t \rightarrow 0} \frac{\langle [u(t + \Delta t) - u_0]^2 | [u_0, t_0] \rangle}{\Delta t} = 1 \quad (76)$$

Die FPG ist dann

$$\frac{\partial}{\partial t} p(u, t | u_0, t_0) = \frac{1}{2} \frac{\partial^2}{\partial u^2} p(u, t | u_0, t_0); \quad (77)$$

hier ist die Drift gleich Null und der Diffusionskoeffizient gleich 1. Der Vergleich mit (43) zeigt, dass diese Gleichung die FPG für den Wiener-Prozess ist. Der durch (71) definierte stochastische Prozess ist also gerade der Wiener-Prozess. Diesem Resultat haftet etwas Paradoxes an: $u(t)$ ist als Integral der Trajektorie $\xi(t)$ eines Prozesses definiert, der als Weißes Rauschen deklariert wurde, erweist sich dann als Wiener-Prozess; die Derivierte von $u(t)$ müßte also wieder $\xi(t)$ sein. Da aber $u(t) = W(t)$, folgt, dass $u(t)$ nirgends differenzierbar ist, denn diese Eigenschaft wurde ja bereits für die Trajektorien von $W(t)$ nachgewiesen. Daraus folgt, streng genommen, dass die Dgl (69), ebenso wie (66), nicht existiert. Andererseits hat aber die Integralgleichung

$$x(t) - x(0) = \int_0^t a[x(s), s] ds + \int_0^t b[x(s), s] \xi(s) ds \quad (78)$$

einen interpretierbaren Sinn. Man definiert nun

$$dW(t) = W(t + dt) - W(t) = \xi(t) dt. \quad (79)$$

Das zweite Integral in (78) kann dann in der Form

$$\int_0^t b[x(s), s]dW(s) \quad (80)$$

geschrieben werden. Das Integral erweist sich damit als eine Art stochastisches Stieltjes Integral³ in bezug auf die Trajektorie $W(t)$. Die Definition muß allerdings noch präzisiert werden.

1.9.2 Definition des stochastischen Integrals

Es sei $G(t)$ eine beliebige Funktion und $W(t)$ ein Wiener-Prozess. Es soll das Integral

$$\int_0^t G(s)dW(s) \quad (81)$$

als eine Art Riemann-Integral erklärt (definiert) werden. Das Intervall $[t_0, t]$ wird dazu in n Teilintervalle $[t_i, t_{i+1})$ aufgeteilt, mit $t_0 \leq t_1 \leq \dots \leq t_{n-1} \leq t$. Weiter werden Zwischenpunkte τ_i mit $t_i \leq \tau_i \leq t_{i+1}$ eingeführt. Das Integral (81) wird dann als Grenzwert der Summe

$$S_n = \sum_{i=1}^n G(\tau_i)[W(t_i) - W(t_{i-1})] \quad (82)$$

definiert.

Es zeigt sich, dass die Summe S_n von der Wahl der Zwischenpunkte τ_i abhängt. Dazu werde $G(\tau_i) = W(\tau_i)$ gesetzt. Dann ist

$$\begin{aligned} \langle S_n \rangle &= \left\langle \sum_{i=1}^n W(\tau_i)[W(t_i) - W(t_{i-1})] \right\rangle \\ &= \sum_{i=1}^n [\min(\tau_i, t_i) - \min(\tau_i, t_{i-1})] \\ &= \sum_{i=1}^n (\tau_i - t_{i-1}). \end{aligned} \quad (83)$$

Setzt man zB $\tau_i = at_i + (1-a)t_{i-1}$ mit $0 < a < 1$, so ist

$$\langle S_n \rangle = \sum_{i=1}^n (t_i - t_{i-1})a = (t - t_0)a.$$

$\langle S_n \rangle$ kann also jeden Wert zwischen 0 und 1 annehmen.

Definition 1.1 Es sei $a = 0$, dh $\tau_i = t_{i-1}$. Dann heißt

$$\int_{t_0}^t G(s)dW(s) = \text{ms} - \lim_{n \rightarrow \infty} \left[\sum_{i=1}^n G(t_{i-1})[W(t_i) - W(t_{i-1})] \right] \quad (84)$$

das *stochastische Ito-Integral*.

³Das gewöhnliche (Riemann-) Integral $\int f(x)dx$ ist der Grenzwert der Summe $\sum_i f(x_i)(x_i - x_{i-1})$. Sei $g(x)$ eine Funktion von x . Dann heißt $\int f(x)dg(x) = \lim \sum_i f(x_i)[g(x_i) - g(x_{i-1})]$ das Stieltjes-Integral von f bezüglich g . Man macht in der Statistik bei der Definition von Erwartungswerten vom Stieltjes-Integral Gebrauch: Es ist $\langle x \rangle = \int xf(x)dx$. Aber $f(x) = dF(X)/dx$, die Dichte f ist die Ableitung der Verteilungsfunktion F . Also ist $dF(x) = f(x)dx$ und man kann $\langle x \rangle = \int x dF(x)$ schreiben. Ist h eine Funktion von x , so hat man $\langle h(x) \rangle = \int h(x)dF(x)$. Diese Integrale sind Stieltjes- oder auch Riemann-Stieltjes-Integrale.

Anmerkung: ms-lim steht für *mean square limit*.

Beispiel 1.1 Es werde

$$\int_{t_0}^t W(s)dW(s)$$

betrachtet. Mit $dW_t = dW(t)$ hat man

$$\begin{aligned} S_n &= \sum_{i=1}^n W_{i-1}(W_i - W_{i-1}) \equiv \sum_{i=1}^n W_{i-1}\Delta W_i \\ &= \frac{1}{2} \sum_{i=1}^n [(W_{i-1} + \Delta W_i)^2 - \Delta W_i^2 - (W_{i-1})^2 - (\Delta W_i)^2] \\ &= \frac{1}{2} [W(t)^2 - W(t_0)^2] - \frac{1}{2} \sum_{i=1}^n (\Delta W_i)^2 \end{aligned} \quad (85)$$

Es kann der *mean square limit* des letzten Terms berechnet werden.

$$\langle \sum_{i=1}^n (\Delta W_i)^2 \rangle = \sum_i \langle (W_i - W_{i-1})^2 \rangle = \sum_i (t_i - t_{i-1}) = t - t_0.$$

Weiter

$$\begin{aligned} \langle [\sum_i (W_i - W_{i-1})^2 - (t - t_0)]^2 \rangle &= \langle \sum_i (W_i - W_{i-1})^4 + 2 \sum_{i>j} (W_i - W_{i-1})^2 (W_j - W_{j-1})^2 \\ &\quad - 2(t - t_0) \sum_i (W_i - W_{i-1})^2 + (t - t_0)^2 \rangle. \end{aligned}$$

$W_i - W_{i-1}$ ist (i) eine Gauss-Variable und (ii) unabhängig von $W_j - W_{j-1}$. Deshalb gilt

$$\langle (W_i - W_{i-1})^2 (W_j - W_{j-1})^2 \rangle = (t_i - t_{i-1})(t_j - t_{j-1})$$

und so weiter, vergl. Gardiner, p. 85.

Man findet

$$\int_{t_0}^t W(s)dW(s) = \frac{1}{2} [W(t)^2 - W(t_0)^2 - (t - t_0)^2]. \quad (86)$$

Dieses Integral ist nicht gleich dem gewöhnlichen Riemann-Stieltjes Integral, bei dem der Term $(t - t_0)^2$ nicht auftreten würde. Der Grund dafür ist, dass $|W(t + \Delta t) - W(t)|$ fast immer von der Ordnung $\sqrt{\Delta t}$ ist. Daraus folgt, dass Terme zweiter Ordnung in $\Delta W(t)$ nicht wie bei der gewöhnlichen Integration verschwinden.

1.9.3 Stochastische Differentialgleichungen

Betrachtet werde noch einmal die Differentialgleichung (69), dh

$$\frac{dX(t)}{dt} = a(X, t) + b(X, t)\xi(t).$$

Die Interpretation dieser Gleichung ist die Integralgleichung

$$x(t) - x(0) = \int_0^t a[x(s), s]ds + \int_0^t dW(s)b[x(s), s]ds. \quad (87)$$

Dieser Sachverhalt führt zu der folgenden

Definition 1.2 Die stochastische Größe $x(t)$ genügt einer Ito stochastischen Differentialgleichung

$$dx(t) = a[x(t), t]dt + b[x(t), t]dW_t \quad (88)$$

dann, wenn für alle t und t_0

$$x(t) = x(t_0) + \int_{t_0}^t a[x(s), s]ds + \int_{t_0}^t b[x(s), s]dW_s \quad (89)$$

gilt.

$x(t)$ ist die Trajektorie eines Markov-Prozesses. Also existiert eine bedingte Dichte $p(x, t|x_0, t_0)$. Es läßt sich zeigen (Gardiner, p. 96), dass diese Dichte der Gleichung

$$\partial_t p(x, t|x_0, t_0) = -\partial_x [a(x, t)p(x, t|x_0, t_0)] + \frac{1}{2} \partial_x^2 [b(x, t)^2 p(x, t|x_0, t_0)] \quad (90)$$

genügt. Damit ist $x(t)$ die Trajektorie eines Diffusionsprozesses mit der Drift $a(x, t)$ und der Diffusion $b(x, t)^2$.

1.10 Beispiele

1.10.1 Drift und Diffusion unabhängig von x

Gegeben sei die Gleichung

$$dx(t) = a(t)dt + b(t)dW_t. \quad (91)$$

Weder der Drift- noch der Diffusionskoeffizient hängen von x ab, dh sie sind nicht-zufällige Funktionen der Zeit. Die Lösung ist einfach

$$x(t) = x_0 + \int_{t_0}^t a(s)ds + \int_{t_0}^t b(s)dW_s \quad (92)$$

Dabei kann x_0 zufällig sein oder nicht. $x(t)$ ist Gauss-verteilt, wenn x_0 deterministisch oder selbst Gauss-verteilt ist, denn das Integral rechts ist eine Linearkombination von infinitesimalen Gauss-Variablen. Für den Erwartungswert erhält man

$$\langle x(t) \rangle = \langle x_0 \rangle + \int_{t_0}^t a(s)ds, \quad (93)$$

und für die Autokorrelation erhält man

$$\begin{aligned} \langle [x(t) - \langle x(t) \rangle][x(s) - \langle x(s) \rangle] \rangle &\equiv \langle x(t), x(s) \rangle \\ &= \left\langle \int_{t_0}^t b(s)dW_s \int_{t_0}^s b(s')dW_{s'} \right\rangle \\ &= \int_{t_0}^{\min(t,s)} [b(s)]^2 ds \end{aligned} \quad (94)$$

Häufig nimmt man an, die Aktivität eines sensorischen "Kanals" sei durch $x(t) = g(t) + \xi(t)$ gegeben, wobei $\xi(t)$ "Rauschen" bedeutet, und $E[x(t)] = \langle x(t) \rangle = g(t)$. Wendet man (93) darauf an, so erhält man

$$g(t) = \langle x_0 \rangle + \int_{t_0}^t a(s)ds, \quad \langle x_0 \rangle = 0$$

und

$$\frac{dg(t)}{dt} = a(t)$$

ist. Wird der Kanal als lineares System aufgefasst, so ergibt sich g als Faltung

$$g(t) = \int_0^\infty \sigma(s)h(t-s)ds,$$

wobei σ das Eingangssignal (Stimulus) ist und h die Impulsantwort. Für $\sigma(t) = c$, $0 \leq t \leq t_s$ erhält man insbesondere

$$g(t) = \begin{cases} c \int_0^t h(t-s)ds, & t \leq t_s \\ c \int_0^{t_s} h(t_s-s)ds, & t > t_s \end{cases}$$

Dann ist $dg/dt = a(t) = ch(t)$ für $t \leq t_s$ und $dg/dt = a(t) = h(t) - h(t-t_s)$ für $t > t_s$.

1.10.2 Das Integrate-and-fire-Modell.

Es wird angenommen, dass ein Neuron ein Aktionspotential ("spike") erzeugt, wenn das Membranpotential u einen Schwellenwert ϑ von unten erreicht. Die Membran wird dabei als eine Parallelschaltung eines Widerstands R und eines Kondensators C mit der Kapazität q aufgefasst. Der Strom $I(t)$ durch die Membran ist dann durch $I(t) = I_R(t) + I_C(t)$ definiert. Nach dem Ohmschen Gesetz ist $I_R = u/R$, während $I_C = C du/dt$ ist, so dass $I_C = C du/dt$. Dann soll

$$I(t) = \frac{u(t)}{R} + C \frac{du}{dt}. \quad (95)$$

Das Produkt $\tau_m = RC$ heißt Zeitkonstante des Systems. Multipliziert man (95) mit R , so erhält man

$$RI(t) = u(t) + RC \frac{du}{dt}, \quad \text{dh } \tau_m \frac{du}{dt} = -u(t) + RI(t). \quad (96)$$

Die *firing time* des Neurons ist $t^{(f)}$, mit

$$t^{(f)} : u(t^{(f)}) = \vartheta. \quad (97)$$

Unmittelbar nach $t^{(f)}$ wird das Membranpotential auf einen neuen Wert $u_r < \vartheta$ zurückgesetzt:

$$\lim_{t \rightarrow t^{(f)}; t > t^{(f)}} u(t) = u_r. \quad (98)$$

Für $t > t^{(f)}$ beginnt die Dynamik dann wieder mit (96).

Bis hierher ist die Dynamik des Neurons deterministisch. Man findet aber eine unregelmäßige Aktivität. Deswegen addiert man einen Term $\xi(t)$, er Rauschen repräsentiert. Man nimmt an, dass $\xi(t)$ Gaußsches Weißes Rauschen repräsentiert, mit $\langle \xi(t) \rangle = 0$ und der Autokorrelation

$$\langle \xi(t)\xi(t') \rangle = \sigma^2 \tau_m (t - t'). \quad (99)$$

Dabei ist σ die Amplitude des Rauschens. Dann erhält man die stochastische Differentialgleichung

$$\tau_m \frac{du}{dt} = -u(t) + RI(t) + \xi(t). \quad (100)$$

Diese Gleichung spezifiziert einen *Ornstein-Uhlenbeck-Prozess*. (100) heißt auch Langevin-Gleichung des "noisy" integrate-and-fire-Modells (vergl. (68)).

Das Neuron habe zur Zeit \hat{t} zuletzt gefeuert. Wegen der zufällig eintreffenden Spikes kann das Potential zur Zeit t nicht vorausberechnet werden. Aber die Wahrscheinlichkeitsdichte $p(u, t)$

kann berechnet werden. Der Einfachheit halber werde $I^{ext} = 0$ gesetzt. Die Input-Spikes an der k -ten Synapse werden durch einen Poisson-Prozess mit der Rate $\nu_k(t)$ erzeugt⁴. Die Wahrscheinlichkeit, dass kein Spike in Δt ankommt, ist

$$P(\text{kein Spike in } [t, t + \Delta t]) = 1 - \sum_k \nu_k(t). \quad (101)$$

Wenn kein Spike während $[t, t + \Delta t]$ am Neuron ankommt, ändert sich das Potential von $u(t) = u'$ zu $u(t + \Delta t) = u' \exp(-\Delta t / \tau_m)$. Kommt ein Spike in diesem Intervall an, ist die Veränderung von u' zu $u' \exp(-\Delta t / \tau_m) + w_k$. Ist also $u(t) = u'$, so ist die W-Dichte für u zur Zeit $t + \Delta t$ durch

$$\begin{aligned} P^{trans}(u, t + \Delta t | u', t) &= \left[1 - \Delta t \sum_k \nu_k(t) \right] \delta(u - u' e^{-\Delta t / \tau_m}) \\ &\quad + \Delta t \sum_k \nu_k(t) \delta(u - u' e^{-\Delta t / \tau_m} - w_k) \end{aligned} \quad (102)$$

u ist ein Markov-Prozess, und es gilt dann

$$p(u, t + \Delta t) = \int P^{trans}(u, t + \Delta t | u', t) p(u', t) du'. \quad (103)$$

Will man das Integral berechnen, muß man die Regel $\delta(at) = a^{-1} \delta(t)$ berücksichtigen. Das Resultat der Integration ist

$$\begin{aligned} p(u, t + \Delta t) &= \left[1 - \Delta t \sum_k \nu_k(t) \right] e^{\Delta t / \tau_m} p(e^{\Delta t / \tau_m} u, t) \\ &= + \Delta t \sum_k \nu_k(t) e^{\Delta t / \tau_m} p(e^{\Delta t / \tau_m} u - w_k, t) \end{aligned} \quad (104)$$

Man kann nun den Fall betrachten, dass Δt klein ist, so dass $p(u, t + \Delta t)$ um $\Delta t = 0$ in eine Reihe entwickelt werden kann. Für die Annäherung erster Ordnung findet man dann

$$\begin{aligned} \frac{p(u, t + \Delta t) - p(u, t)}{\Delta t} &= \frac{1}{\tau_m} p(u, t) + \frac{1}{\tau_m} u \frac{\partial}{\partial u} p(u, t) \\ &\quad + \sum_k \nu_k(t) [p(u - w_k, t) - p(u, t)]. \end{aligned} \quad (105)$$

Für $\Delta t \rightarrow 0$ wird die linke Seite zur partiellen Ableitung $\partial p(u, t) / \partial t$. Sind die Sprungamplituden w_k klein, kann man die rechte Seite bezüglich u in eine Reihe entwickeln:

$$\begin{aligned} \tau_m \frac{\partial}{\partial t} p(u, t) &= - \frac{\partial}{\partial u} \left[-u + \tau_m \sum_k \nu_k(t) w_k \right] p(u, t) \\ &\quad + \frac{1}{2} \left[\tau_m \sum_k \nu_k(t) w_k^2 \right] \frac{\partial^2}{\partial u^2} p(u, t). \end{aligned} \quad (106)$$

Dies ist die *Kramers-Moyal-Expansion*; sie ist ein Beispiel einer Fokker-Planck-Gleichung. Der erste Klammerterm ist die Drift des Membranpotentials, die sich durch die "leakage" + den mittleren Hintergrundeingang ($\propto \sum_k \nu_k(t) w_k$) ergibt. Der zweite Ausdruck in Klammern ist der Diffusionsterm, der die Fluktuationen des Membranpotentials spiegelt. (106) ist der Langevin-Gleichung (100) äquivalent, mit $RI(t) = \tau_m \sum_k \nu_k(t) w_k$ und der zeitabhängigen Rauschamplitude

$$\sigma^2(t) = \tau_m \sum_k \nu_k(t) w_k^2. \quad (107)$$

Für den Spezialfall $\sigma^2(t) = \sigma^2$ eine Konstante erhält man den Ornstein-Uhlenbeck-Prozess.

⁴Die Wahrscheinlichkeit des Auftretens von j Spikes ist durch $P(j) = \exp(-\lambda) (\Delta t \lambda)^j / j!$ gegeben.

2 Populationsaktivität

Gegeben seien N Neurone; die Aktivität dieser Population kann mit der Anzahl der aktiven Neurone gleichgesetzt werden. Dazu wird die Anzahl der Spikes $n_a(t, t + \Delta t)$ für kleines Δt gezählt und durch N geteilt werden. Teilt man nur durch Δt , so erhält man die Populationsaktivität

$$A(t) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \frac{n_a(t; t + \Delta t)}{N}. \quad (108)$$

Die Population sei homogen, dh es soll gelten

$$I_i^{ext}(t) = I^{ext}(t), \quad \text{für alle } i \quad (109)$$

$$w_{ij} = \frac{J_0}{N}, \quad (110)$$

wobei i das i -te Neuron ist, und w_{ij} die Kopplung zwischen dem i -ten und dem j -ten Neuron ist; die w_{ij} sollen also alle gleich sein, mit J_0 als Kopplungsparameter. Ist $J_0 = 0$, so sind die Neurone unabhängig voneinander. Für $J_0 > 0$ sind die Verbindungen exzitatorisch, für $J_0 < 0$ sind sie inhibitorisch. Die Neurone sind zB *leaky-integrate-and-fire* Neurone:

$$\tau_m \frac{d}{dt} u_i = -u_i + RI_i(t), \quad i = 1, \dots, N \quad (111)$$

I_i ist durch

$$I_i(t) = \sum_{j=1}^N \sum_f w_{ij} \alpha(t - t_j^{(f)}) + I^{ext}(t) \quad (112)$$

gegeben. Nach dem Ohmschen Gesetz gilt $u = RI$, $RI_i(t)$ ist also eine Spannung und entspricht damit u_i . $\alpha(t - t_j^{(f)})$ entspricht einem postsynaptischen Strom. Der gesamte Input-Strom ist identisch für alle Neurone:

$$I(t) = J_0 \int_0^\infty \alpha(s) A(t - s) ds + I^{ext}(t). \quad (113)$$

Es sei $p(u, t)$ die Dichte des Membranpotentials. Es wird nach dem Anteil der Neurone gefragt, deren Potential zwischen u_0 und $u_0 + \Delta u$ liegt. Man findet

$$\lim_{N \rightarrow \infty} \left[\frac{\text{Neurone mit } u_0 < u_i(t) \leq u_0 + \Delta u}{N} \right] = \int_{u_0}^{u_0 + \Delta u} p(u, t) du. \quad (114)$$

Es gilt

$$\int_{-\infty}^{\vartheta} p(u, t) du = 1, \quad (115)$$

dh die Wahrscheinlichkeit, dass alle Neurone ein Potential kleiner als, höchstens gleich der Schwelle ϑ haben, ist gleich 1. Der Anteil der Neurone, die die Schwelle erreichen - über die Schwelle "fließen" - gleicht dem Erwartungswert der Populationsaktivität. Dieser "Flux" sei $J(\vartheta, t)$; dann gilt also

$$A(t) = J(\vartheta, t). \quad (116)$$

Die Neurone, die über die Schwelle ϑ "verschwinden", werden de facto auf den Wert u_r zurückgesetzt und tauchen deshalb auch gleich wieder in der Neuronenmenge auf.

Alle Neurone erhalten den gleichen treibenden Inputstrom $I(t)$. Darüber hinaus erhalten sie einen stochastischen Input von anderen Neuronen (Hintergrundstimulation). Dieser Input wirkt über Synapsen verschiedenen Typs. Ein Spike an der Synapse vom k -ten Typ erzeugt

einen Sprung des Membranpotentials von der Größe w_k . Die effektive Spikerate, summiert über alle Synapsen vom k -ten Typ, sei ν_k . Es wird angenommen, dass die verschiedenen *spike trains* an verschiedenen Neuronen und verschiedenen Synapsen - stochastisch unabhängig voneinander sind. Damit erhält man für die Dynamik und $u \leq \vartheta$

$$\begin{aligned} \frac{\partial}{\partial t} p(u, t) &= \frac{1}{\tau_m} p(u, t) - \frac{1}{\tau_m} [-u + RI^{ext}(t)] \frac{\partial}{\partial u} p(u, t) \\ &+ \sum_k \nu_k(t) [p(u - w_k, t) - p(u, t)] + A(t) \delta(u - u_r), \end{aligned} \quad (117)$$

und

$$p(u, t) = 0, \quad u > \vartheta, \quad (118)$$

vergl. (105). Um $A(t)$ zu berechnen, muß die Flux durch die Schwelle ϑ bestimmt werden. Dazu wird die Flux $J(u_0, t)$ durch ein beliebiges Referenzpotential u_0 betrachtet:

$$J(u_0, t) = J_{drift}(u_0, t) + J_{jump}(u_0, t), \quad (119)$$

wobei J_{drift} die stetige Drift des Membranpotentials bezeichnet, die stattfindet, wenn kein Spike eintrifft. J_{jump} ergibt sich aus den exzitatorischen und inhibitorischen Spikes.

Es sei $w_k > 0$ der exzitatorische Input. Alle Neurone mit einem Potential u_i , $u_0 - w_k < u_i \leq u_0$ "springen" durch das Referenzpotential u_0 , wenn ein Spike an einer Synapse vom Typ k ankommt. Die zugehörige Spikerate ist ν_k , also ist die gesamt Flux durch

$$J_{jump}(u_0, t) = \sum_k \int_{u_0 - w_k}^{u_0} p(u, t) du. \quad (120)$$

Die Drift $J_{drift}(u_0, t)$ ergibt sich aus der Dichte $p(u_0, t)$ des Potentials u_0 , multipliziert mit der momentanen Veränderung du/dt ,

$$du/dt = [-u + RI^{ext}(t)]/\tau_m,$$

und so erhält man

$$J_{drift}(u_0, t) = \frac{1}{\tau_m} [-u_0 + RI^{ext}(t)] p(u_0, t). \quad (121)$$

Die gesamt Flux an der Schwelle $u_0 = \vartheta$ liefert dann die Populationsaktivität

$$A(t) = \frac{1}{\tau_m} [-\vartheta + RI^{ext}(t)] p(\vartheta, t) + \sum_k \nu_k \int_{\vartheta - w_k}^{\vartheta} p(u, t) du. \quad (122)$$

(117) und (122) definieren zusammen die Dynamik in einer Population von *integrate-and-fire*-Neuronen mit stochastischem Hintergrundinput.

Es werde der Anteil von Neuronen mit einem Membranpotential u , $u_0 < u < u_1$ betrachtet. Der Anteil wächst durch Flux von unten, dh durch u_0 , und von oben, dh durch u_1 . Die Flux $J(u, t) > 0$ bedeutet ein Steigen des Anteils. Man hat den Erhaltungssatz

$$\frac{\partial}{\partial t} \int_{u_0}^{u_1} p(u', t) du' = J(u_0, t) - J(u_1, t). \quad (123)$$

Für $u = \vartheta$ und $u = u_r$ erfolgt eine Rücksetzung des Potentials ("reset"). Nimmt man die Ableitung mit Bezug auf die obere Grenze u_1 und wechselt die Bezeichnung von u_1 in u um, erhält man die Beziehung

$$\frac{\partial}{\partial t} = - \frac{\partial}{\partial u} J(u, t), \quad u \neq u_r, \quad u \neq \vartheta. \quad (124)$$

Neurone, die einen Spike erzeugt haben, beginnen eine neue Trajektorie bei u_r und bilden eine "Quelle" von Trajektorien, dh es erscheinen neue Trajektorien im Intervall $u_r - \epsilon, u_r + \epsilon$, die in das Intervall nicht durch die Grenzen eingehen. Um für diese Trajektorien zu sorgen, wird der Term $A(t)\delta(u - u_r)$ auf der rechten Seite von (124) addiert.

Geht man zu infinitesimalen w_k über, erhält man eine Diffusionsapproximation (Entwicklung von (117) in eine Taylor-Reihe); man erhält

$$\begin{aligned} \tau_m \frac{\partial}{\partial t} p(u, t) = & - \frac{\partial}{\partial u} \left\{ \underbrace{\left[-u + RI^{ext}(t) + \tau_m \sum_k \nu_k(t) w_k \right]}_{Drift} p(u, t) \right\} \\ & + \underbrace{\frac{1}{2} \frac{\partial^2}{\partial u^2} \left[\tau_m \sum_k \nu_k(t) w_k^2 p(u, t) \right]}_{Diffusion} + \tau_m A(t) \delta(u - u_r) + O(w_k^3). \end{aligned} \quad (125)$$

Der Diffusionsterm hängt hier *nicht* von u ab, definiert aber das durch die Hintergrundsaktivität erzeugt Rauschen, dass damit als additiv angenommen wird. Andererseits repräsentieren die λ_k Poisson-Parameter verschiedenen Typs, also Rauschen verschiedener Intensität, dass dann allerdings zu einem Rauschprozess zusammengefasst wird. Hier wäre es möglich, einige der Poisson-Prozesse mit konstantem Parameter $\lambda_k(t) = \lambda_k$ "laufen" zu lassen, andere aber zeitlich zu modulieren, so dass eine langsam fluktuierende Komponente eingebaut würde.

Die Schwelle ϑ ist eine absorbierende Grenze, so dass

$$p(\vartheta, t) = 0. \quad (126)$$

Für die Flux durch die Schwelle ergibt sich durch Reihenentwicklung von (122)

$$A(t) = - \frac{\sigma^2(t)}{2\tau_m} \frac{\partial p(u, t)}{\partial u} \Big|_{u=\vartheta}, \quad \sigma^2(t) = \tau_m \sum_k \nu_k(t) w_k^2. \quad (127)$$