LECTURE 2-3: Stochasticity

PDF available at http://math.la.asu.edu/~chris/CIME09/CIME09.htm
Part 1: Traffic flow vs. production systems

Part 2: First principle models.
- Simple deterministic first principle models → hyperbolic conservation laws.
- Stochasticity and kinetic models.
- Mean field theory, long time averages and diffusive corrections.

Part 3: Clearing distributions from observed data.

Part 4:
- The problem with diffusive corrections due to stochastic fluctuations. Hyperbolic relaxation models.
Part 1: Traffic flow models (fluid dynamics)

- Introduce an artificial variable \( x \), the stage of the production process. Parts enter at stage \( x = 0 \). Finished product leaves at stage \( x = 1 \). Parts travel on the ’road’ from \( x = 0 \rightarrow x = 1 \).

- Traffic models for production networks constitute an intermediate stage between DES and rate equations. PDE’s \( \Rightarrow \) amenable to analysis and optimization.

- Large body of theory: Lighthill, Whitham, Prigogine, Daganzo, Helbing, Peters, Klar, Rascle
Production Systems vs. Traffic Flow

Conservation Laws:

$$\partial_t \rho + \partial_x [\phi(\rho)] = 0$$

- \(\rho\): Work in progress (WIP-) density or density of vehicles
- \(x\): stage of the production process. \(x = 0\): unfinished product,
- \(x = X\): finished product, traffic flow: road
- \(\phi\): flux function
Similarities between Traffic and Production Models

- **Complexity and Topology:** Complex re-entrant production systems. Networks of roads.

- **Control:** Policies for production systems. Traffic control mechanisms.

- **Random behavior.**

- **First Principle Models:** Discrete Event Simulation (DES), Multi-Agent Models (incorporate stochastic behavior) $\Rightarrow$ kinetic equations for densities (mean field theories, large time asymptotics) $\Rightarrow$ fluid dynamics $\Rightarrow$ rate equations (fluid models).

- **Simulation** $\Rightarrow$ optimization and control.
Differences between Traffic and Production Modeling

- **Capacity Limits:** Limited capacity of processors (machines), limited space (capacity of the road) \(\Rightarrow\) backward wave propagation in traffic flow \(\frac{\phi(\rho)}{\rho} > 0\) and \(\phi'(\rho) < 0\).

- **Parameters and Control Variables:**
  - Traffic: randomly given influx and individual behavior, distribute road capacities.
  - Production: randomly given demand, choose start rate and (to some extent) topology.
Arrivals and processing times governed by Markov processes:

\[
v(x, \rho) = \frac{c(x)}{1 + \rho}, \quad c(x) = \frac{1}{\langle \text{processing times} \rangle}
\]

\(c(x)\): service rate or capacity of the processor at stage \(x\).

Simplest traffic flow model (Lighthill - Whitham - Richards)

\[
v(x, \rho) = v_0(x) \left(1 - \frac{\rho}{\rho_{\text{jam}}} \right)
\]

- In supply chain models the density \(\rho\) can become arbitrarily large, whereas in traffic the density is limited by the space on the road \(\rho_{\text{jam}}\).
The phase velocity $v_{\text{phase}}$

$$v_{\text{phase}} = \frac{\partial}{\partial \rho} [\rho v(x, \rho)]$$

$$v_{\text{phase}} = \frac{c(x)}{1 + \rho^2} > 0, \quad v_{\text{phase-traffic}} = v_0(x) \left[1 - \frac{2\rho}{\rho_{\text{max}}} \right]$$

In supply chain models the propagation of information (shock speeds) is strictly forward $v_{\text{phase}} > 0$, whereas in traffic flow models shock speeds can have both signs.
Problem:

- Queuing theory models are based on quasi-steady state regime. Modern production systems are almost never in steady state. (short product cycles, just in time production).

- Goal: Derive non-equilibrium models from first principles (first for automata) and then including stochastic effects.
CONTENTS

- Part 1: Traffic flow vs. production systems
- Part 2: First principle models.
  - Simple deterministic first principle models $\rightarrow$ hyperbolic conservation laws.
  - Stochasticity and kinetic models.
  - Mean field theory, long time averages and diffusive corrections.
- Part 3: Clearing distributions from observed data.
- Part 4:
  - The problem with diffusive corrections due to stochastic fluctuations. Hyperbolic relaxation models.
Assume processors work deterministically like automata. A processor located in the infinitesimal stage interval of length $\Delta x$ needs a time $\tau(x) = \frac{\Delta x}{v_0(x)}$ to process an item.

It cannot accept more than $c(x) \Delta t$ items per infinitesimal time interval $\Delta t$. 
N-curves (Newell 1958):

**Goal:** Find flux function for \( \partial_t \rho + \partial_x \phi = 0 \)

\[
U(x, t) = \int_{-\infty}^{t} \phi(x, s) \, ds \implies \rho(x, t) = -\partial_x U
\]

Model the inverse of \( U \) w.r.t time:

\[
U(x, t) = s \iff Z(x, s) = t
\]

\( U(x, t) \): Number of parts processed by processor \( x \) at time \( t \)

\( Z(x, s) \): The time processor \( x \) processes part number \( s \).
Modeling $Z$

- $T_a(x, s)$: arrival time at beginning of queue,
- $T_f$: time part $s$ is fed into processor $x$

\[
T_f(x, s) = \max\{T_a(x, s), T_f(x, s - \Delta s) + \frac{\Delta s}{c(x)}\}
\]

\[
T_a(x + \Delta x, s) = T_f(x, s) + \frac{\Delta x}{v_0(x)} = Z(x, s)
\]

\[
Z(x, s) = \max\{Z(x - \Delta x, s) + \frac{\Delta x}{v_0(x)}, Z(x, s - \Delta s) + \frac{\Delta s}{c(x)}\}
\]
The inverse relations

\[ Z(s, x) = t \iff U(x, t) = s \]

\[ \Rightarrow \partial_s Z = \frac{1}{\partial_t U}, \quad \partial_x Z = -\frac{\partial_x U}{\partial_t U} \]

\[ \Rightarrow 0 = \max\{\Delta x \frac{\partial x U}{\partial t U} + \frac{\Delta x}{v_0(x)}, -\Delta s \frac{1}{\partial_t U} + \frac{\Delta s}{c(x)}\} \]

\[ \Rightarrow \partial_t U = \min\{-v_0 \partial_x U, c\} \]

\[ \partial_t U = \phi, \quad -\partial_x U = \rho, \quad \partial_t \rho + \partial_x \phi = 0 \]

\[ \Rightarrow \partial_t \rho + \partial_x [\min\{c(x), v_0(x) \rho\}] \]
The simple deterministic automaton model yields the conservation law

\[ \Rightarrow \partial_t \rho + \partial_x \left[ \min\{c(x), v_0(x)\rho\} \right] \]
\[ \partial_t \rho + \partial_x \phi(x, \rho) = 0, \quad \phi(x, \rho) = \min\{c(x), v_0(x)\rho\} \]

- No maximum principle (similar to pedestrian traffic with obstacles).
- The capacity \( c(x) \) is discontinuous if nodes in the chain form a bottleneck.
- Flux \( \phi \) discontinuous \( \Rightarrow \) density \( \rho \) distributional. (alternative model by Klar, Herty ’04).
- Random server shutdowns \( \Rightarrow \) bottlenecks shift stochastically.
A bottleneck in a continuous supply chain

Temporary overload of the bottleneck located at \( x = 1 \).
Deterministic models produce a linear steady state WIP - flux relation below capacity, and no steady state above capacity.

Stochastic models produce a diagram with the capacity as an asymptote.

⇒ arbitrarily large steady state WIP below capacity.
Flux vs. WIP in steady state for deterministic and stochastic models.
Part 1: Traffic flow vs. production systems

Part 2: First principle models.
  • Simple deterministic first principle models $\rightarrow$ hyperbolic conservation laws.
  • Stochasticity and kinetic models.
  • Mean field theory, long time averages and diffusive corrections.

Part 3: Clearing distributions from observed data.

Part 4:
  • The problem with diffusive corrections due to stochastic fluctuations. Hyperbolic relaxation models.
Stochastic effects:

- Random arrivals (external influx), random breakdown of servers (medium).
- Random effects significantly influence the behavior.
Divide $[0, 1]$ into $K$ cells, each corresponding to one processor.

$$c(x, t) = \sum_k \chi_k(x) \mu_k(t)$$

The service rates $\mu_k(t)$ switch between $\mu_k = \mu_k^{up}(x)$ and $\mu_k = 0$ according to a Markov process.

\[
\mu_k(t + \Delta t) = \begin{pmatrix}
\mu_k(t) & \text{prob} = 1 - \Delta t \frac{\omega_k(\mu_k)}{\varepsilon} \\
\mu_k^{up} - \mu_k(t) & \text{prob} = \Delta t \frac{\omega_k(\mu_k)}{\varepsilon}
\end{pmatrix}
\]
\[ \mu_k(t + \Delta t) = \begin{pmatrix} \mu_k(t) & \text{prob} = 1 - \Delta t \frac{\omega_k(\mu_k)}{\varepsilon} \\ \mu_k^{up} - \mu_k(t) & \text{prob} = \Delta t \frac{\omega_k(\mu_k)}{\varepsilon} \end{pmatrix} \]

- \( \frac{\omega_k}{\varepsilon} \) is the switching frequency \( \gg 1 \).

- This is a Markov process, since the probability to change the state is independent of the time elapsed since the last change.
One realization of one $\mu_k$. 

\begin{center}
\includegraphics[width=\textwidth]{sample_capacity_mu.png}
\end{center}
One realization of the density $\rho$

with flux $\phi = \min\{\mu, v\rho\}$ using a stochastic $\mu$
Stochastic fluctuations in conservation laws

- Fluctuations cannot (in general) be included directly on the level of conservation laws.
- Include randomness in an underlying kinetic model.
- The kinetic model is very high dimensional (many body problem).
- Involves the motion of $N$ parts in a medium governed by $K$ random variables (the state of the processors).
Consider $N$ part(icle)s moving through state space $x = [0, 1]$. Part(icle) number $n$ moves with velocity $v_n$ and interacts with the medium (the production system). The production system is described by a state $C = (c_1, \ldots, c_K)$. The state of the production system changes randomly.

\[
x_n(t + \Delta t) = x_n(t) + \Delta t v_n(X, C), \quad d\mathcal{P}[q_k = p] = u_k(p, C) \, dp
\]

\[
c_k(t + \Delta t) = \begin{pmatrix} c_k(t) &\mathcal{P} = 1 - \Delta t \omega_k(C) \\ q_k &\mathcal{P} = \Delta t \omega_k(C) \end{pmatrix}
\]
The many body problem

\[ x_n(t+\Delta t) = x_n(t) + \Delta t v_n(X, C), \quad \mu_k(t+\Delta t) = \begin{pmatrix} c_k(t) \\ q_k \end{pmatrix} \begin{pmatrix} P = 1 - \Delta t \\ P = \Delta t \end{pmatrix} \]

\[ d\mathcal{P}[q_k = p] = u_k(p, C) \, dp \]

\[ \omega_k \text{ Define: } F(X, C, t) \text{ probability density that } (x_1, \ldots, x_N)(t) = X \]

and, at the same time, \((c_1, \ldots, c_K)(t) = C\).
\[ \partial_t F + \nabla_X \cdot (VF) = \frac{1}{\varepsilon} Q[F], \]

\[ Q[F] = \sum_k \left[ \int u_k(c_k, C'_k) \omega_k(C'_k) F(X, C'_k, t) \, dc'_k - \omega_k(C)F \right] \]

\[ C'_k = (c_1, \ldots, c'_k, \ldots, c_K), \quad V = (v_1, \ldots, v_N), \]

This is a generalized version of the classical many body problem.

\( F(X, C, t) \) is a function of \( N + K + 1 \) variables \( \Rightarrow \) Contains all the information about correlations up to any order, but not feasible for computations.
General Approach

- Mean field theory: replace many body equation for $N$ particles by an effective one-body equation for a single particle.

- Large time limit for rapid interactions.

\[
\partial_t F + \nabla_X \cdot (VF) = \frac{1}{\varepsilon} Q[F],
\]

\[
Q[F] = \sum_k \left[ \int u_k(c_k, c'_k) \kappa_k(C'_k) F(X, C'_k, t) \, dc'_k - \kappa_k(C') F \right]
\]
General Idea:

\[ \partial_t F(x_1, \ldots, x_N, t) + \nabla_x (V(X)F) = 0 \]

Assume many identical and independent particles \((N \gg 1)\).

\[ F(x_1, \ldots, x_N, t) \approx \prod_{1}^{N} f(x_n, t) \]

\[ \partial_t f(x_1, t) + \nabla_{x_1} (\tilde{V}[f]f) = 0, \quad \tilde{V}[f] = \int V(X) \prod_{n=2}^{N} f(x_n, t) \, dx_2. \]
The Lagrangian picture:

\[ V_n(X, t) = \min\left\{ \frac{c(x_n)}{\rho(x_n)}, \sum_j \chi_j(x_n)v_0(x_n) \right\} \]

Specific volume:

\[ \frac{1}{\rho(x_n)} = \min\{x_j - x_n : x_j > x_n\} \]
Modification:

\[ \partial_t F(x_1, \ldots, x_N, t) + \nabla_X (V(X)F) = 0 \]

Particles independent of each other but not of the medium. \( \Rightarrow \)
statistical independence for the conditional probability.

\[
F(X, C, t) = \frac{F_0(X, C, t)}{G(C, t)}, \quad G(C, t) = \int G(X', C, t) \, dX'
\]

\[
F_0(X, C, t) = \prod_n f(x_n, C, t)
\]

\[
\partial_t f(x_1, C, t) + \frac{1}{G} \nabla_{x_1} (\tilde{V}[f] G f(x, C)) = 0
\]
Theorem (Degond, CR ’06)

For $N \to \infty$ the mean field velocity $\tilde{V}$ is given by

$$\tilde{V}[f](x, C, t) = \frac{\mu(x, C)}{G} \left[1 - \exp\left(-\frac{v_0(x)fG}{\mu(x, C)}\right)\right],$$

$$\mu(x, C) = \sum_k \chi_k(x)c_k$$

$$\Rightarrow \partial_t(fG) + \nabla_x(\tilde{V}fG) = \frac{1}{\varepsilon}Q[fG]$$
Large time asymptotics

\[
\partial_t (fG) + \nabla_x (\tilde{V} f G) = \frac{1}{\varepsilon} Q[fG]
\]

(\varepsilon \to 0) to derive an equation for

\[
\rho(x, t) = \int f(x, C, t) G(C, t) \, dC
\]
\[
\frac{\partial_t}{\partial t}(fG) + \nabla_x (\tilde{V} fG) = \frac{1}{\varepsilon} Q[fG]
\]

set \( g(x, C, t) = fG \)

\[
\frac{\partial_t}{\partial t}g(x, C, t) + \partial_x [\tilde{V}(x, C)g] = \frac{1}{\varepsilon} Q[g]
\]

Separate the time scales by projecting onto the kernel of \( Q \) and its orthogonal complement.

\[
Q[g] = 0 \iff g(x, C, t) = \rho(x, t) G_0(x, C), \quad \int G_0 \, dC = 1, \, \forall x
\]

\[
P g(x, C, t) = \rho(x, t) G_0(x, C), \quad \int (I - P) g(x, C, t) \, dC = 0
\]
\[ Pg(x, C, t) = \rho(x, t) G_0(x, C), \quad \int (I - P) g(x, C, t) \, dC = 0 \]

\[ \Rightarrow PQ = QP = 0 \]

Separate the time scales:

\[ g(x, C, t) = g_0 + \varepsilon g_1, \quad g_0 = Pg, \quad \varepsilon g_1 = (I - P)g \]

\[ (1) \quad g_0 + P \partial_x [\tilde{V}(g_0 + \varepsilon g_1)] = 0, \]

\[ (2) \quad \varepsilon \partial_t g_1 + (I - P) \partial_x [\tilde{V}(g_0 + \varepsilon g_1)] = Q[g_1] \]
\[ (1) \quad g_0 + P \partial_x [\tilde{V}(g_0 + \varepsilon g_1)] = 0, \]

\[ (2) \quad \varepsilon \partial_t g_1 + (I - P) \partial_x [\tilde{V}(g_0 + \varepsilon g_1)] = Q[g_1] \]

(1) gives the macroscopic equation on the slow time scale.

\[ g_0 = \rho G_0 \Rightarrow \partial_t \rho + \partial_x \left[ \int \tilde{V}(g_0 + \varepsilon g_1) \, dC \right] = 0 \]

(2) gives the infinite dimensional remainder on the fast time scale.
\[ (2) \varepsilon \partial_t g_1 + (I - P) \partial_x [\tilde{V} (g_0 + \varepsilon g_1)] = Q[g_1] \]

Solve (2) asymptotically for \( \varepsilon \to 0 \)

\[ (2) \to (I - P) \partial_x [\tilde{V} g_0] = Q[g_1] \]

This gives the macroscopic system

\[
\begin{align*}
\partial_t \rho (x, t) + \partial_x \left[ \int \tilde{V} (\rho G_0 + \varepsilon g_1) \, dC \right] &= 0 \\
(I - P) \partial_x [\tilde{V} \rho G_0] &= Q[g_1]
\end{align*}
\]

The term \( \varepsilon \int \tilde{V} g_1 \, dC \) gives a small diffusive correction to the hyperbolic equation.
The zero order term

\[ g = f G(x, C, t) = \rho(x, t)G_0(C), \quad Q[G_0] = 0 \]

\[ \Rightarrow \rho(x, t) \text{ satisfies the expectation of the LHS under the equilibrium measure } G_0. \]

\[ \partial_t \rho + \nabla_x (\tilde{v} \rho) = 0, \quad \tilde{v} = \frac{a \mu_{up}}{\rho} [1 - \exp\left(-\frac{v_0(x)\rho}{\mu_{up}(x)}\right)] \]

\(a\): availability \(= \sum_k \chi_k(x) \frac{\omega_k(0)}{\omega_k(0) + \omega_k(\mu_{up})})\)
\[ \tilde{v} = \frac{a \mu^{up}}{\rho} [1 - \exp\left(-\frac{v_0(x) \rho}{\mu^{up}(x)}\right)] \]

- Compare \( \phi \) to the expectation flux \( \phi_E \).

- \( \mu \) is replaced by the 'on' capacity \( \mu^{up} \).

- The function \( \min\{\mu, v\rho\} \) is replaced by \( \mu [1 - e^{-\frac{v\rho}{\mu}}] \), which has the same asymptotic properties for \( \rho \to 0 \) and \( \rho \to \infty \).

- The whole flux \( \phi \) is multiplied by the factor \( \tau_{up}^{av} = \frac{\langle \tau_{up} \rangle}{\langle \tau_{up} \rangle + \langle \tau_{down} \rangle} \).
Flux diagram 51

deterministic vs. $M/M/1$ queue vs. long time average mean field theory
Analysis and model based on Markov processes.
Not appropriate for repair and scheduled maintenance.
The zero order term only uses information about the expectations of the process.
Stochastic fluctuations influence the higher order terms in the expansion.
Fluctuations

The $O(\varepsilon)$ term in the expansion:

$$g(x, C, t) = fG = \rho(x, t)G_0(C) + \varepsilon g_1(x, C, t)$$

$$\partial_t \rho(t, x) + \partial_x \{ ac_{up}(x)[1 - \exp(-\frac{v_0\rho}{c_{up}})] - a\varepsilon \mathcal{V}^2 \partial_x \rho \} = 0$$

$a(x)$: availability $= \frac{\langle T_{up} \rangle}{\langle T_{up} \rangle + \langle T_{down} \rangle}$, ($\langle T \rangle = \frac{1}{\omega}$)

$\mathcal{V} = \frac{\sigma}{\langle T \rangle}$: variation coefficient ($= 1$ for a Markov process).

Since we consider many rapidly switching processors, stochastic fluctuations average out, and appear only as an $O(\varepsilon)$ correction.
Problem:

\[
c(x, t + \Delta t) = \begin{pmatrix}
c(x, t) & \text{prob = } 1 - \Delta t\omega(x, c) \\
c^{up}(x) - c(x, t) & \text{prob = } \Delta t\omega(x, c)
\end{pmatrix}
\]

- No memory of how long the processor has been running or down.
- \(T_{up}, T_{down}\) (times between switches) distributed according to exponential distributions \(\Rightarrow V = 1\).

\[
\mathcal{P}[T_{up/down} = \tau] = \frac{1}{\omega(\mu^{up}/0)} \exp[-\omega\tau]
\]
Analogy to 'intelligent particles' with memory (drivers).

Model an arbitrary stochastic process, by enlarging the state space.
\( \tau \): time elapsed since the last change of state of the processor.

\[
c(x, t + \Delta t) = \begin{cases} 
    c(x, t) & \text{prob} = 1 - \Delta t \omega(x, c, \tau) \\
    c^{up}(x) - c(x, t) & \text{prob} = \Delta t \omega(x, c, \tau)
\end{cases}
\]

\[
\tau(x, t + \Delta t) = \begin{cases} 
    \tau(x, t) + \Delta t & \text{prob} = 1 - \Delta t \omega(x, c, \tau) \\
    0 & \text{prob} = \Delta t \omega(x, c, \tau)
\end{cases}
\]

- Larsen ('07) (Radiative transfer models in clouds)
Lemma

\( u(x, c, s) \, ds \) : distribution of the time between scattering events.

\[
u(x, c, s) \, ds = dP[\tau = s] = \omega(x, c, s) \exp[-\int_0^s \omega(x, c, q) \, dq] \, ds
\]

\( \Rightarrow \) given any kind of probability distribution \( u(x, c, \tau) \, d\tau \) of up / down times, we define a (\( \tau \) dependent) frequency \( \omega \)

\[
\omega(x, c, \tau) = \frac{u(x, c, \tau)}{1 - \int_0^\tau u(x, c, s) \, ds}
\]
\[
u(x, c, s) \, ds = d\mathcal{P}[\tau = s] = \omega(x, c, s) \exp\left[-\int_0^s \omega(x, c, q) \, dq\right] \, ds
\]

This is a generalization of the Markov process, since, for
\[
\omega(x, c, \tau) = \omega(x, c)
\]
we have
\[
u(x, c, \tau) = \omega(x, c) \exp[-\tau \omega(x, c)]
\]
Using this trick, everything stays the same, except for the fact that
the conditional probability (in the mean field assumption) satisfies

\[
\partial_t f(t, x, C, \tau^\rightarrow) + \partial_x \Phi = \frac{1}{\varepsilon} Q[f],
\]

\[
Q[f] = -\nabla_{\tau^\rightarrow} f + \sum_j \delta(\tau_j) \int K(x, C, C') \Omega f(t, x, C', \tau^\rightarrow) \ dC' \tau' - \Omega(x, C, \tau^\rightarrow).
\]
Large time asymptotics (Chapman - Enskog) gives the same result

$$\partial_t \rho(t, x) + \partial_x \{a c_{up}(x)[1 - \exp(-\frac{v_0 \rho}{c_{up}})] - a \varepsilon D(\rho) \mathcal{V}^2 \partial_x \rho\}$$

► $\mathcal{V}$: Variation coefficient for the up and down times of the general underlying switching process.
Numerical Verification - Experiment 1:

Quantitative comparison of steady states

- 40 processors with a peak capacity $c_{up} = 2$ and $v = 1$.
- A bottleneck in processors $11 - 20$ (peak capacity $c_{up} = 1$)
- They are on half of the time
  \[ \langle T_{up} \rangle = \frac{1}{\omega(2)} = \langle T_{down} \rangle = \frac{1}{\omega(0)} = \frac{1}{20} \]
- Solve with the flux $\phi = \min\{c, v\rho\}$ for 200 realizations of $\mu$ and average.
- Constant influx $\phi(x = 0) = 0.25$
- Compare to the solution for $\langle \rho \rangle$ of the equation or the mean and variance.
One realization
Comparison for constant influx $F(x = 0) = 0.25$
Experiment 2:

The transient response
Influx temporarily above the bottleneck capacity $c_{btlck} = 0.5$; average over 500 realizations.
Part 1: Traffic flow vs. production systems

Part 2: First principle models.
  • Simple deterministic first principle models → hyperbolic conservation laws.
  • Stochasticity and kinetic models.
  • Mean field theory, long time averages and diffusive corrections.

Part 3: Clearing distributions from observed data

Part 4:
  • The problem with diffusive corrections due to stochastic fluctuations. Hyperbolic relaxation models.
Part 3: Clearing distributions from observed data

- Observe a specific system (experimentally) for a period of time.

- Compute a (large) table:
  
  \[ a_{mn}, m = 1 : M, \ n = 1 : N \]  
  
  time part number \( n \) has completed stage number \( m \) of the process.

- Extract transport coefficients directly from observed data.

- Build a kinetic model using numerical distributions.

- Construct a macroscopic model via long time averages, using means and variances of the observed distributions.
A kinetic model

\[ x = \xi_n(t) \] stage of part number \( n \) at time \( t \). \( v_n \): velocity

\[ \xi_n(t + \Delta t) = \xi_n(t) + \Delta t v_n(t) \]

\( v_n \) changes randomly, according to a distribution extracted from the observed data.

\[ v_n(t + \Delta t) = (1 - r)v_n(t) + r\eta_n, \quad \mathcal{P}[r = 1] = \omega \Delta t \]

\[ d\mathcal{P}[\eta_n = v] = U(x, t, v) \, dv \]
The kinetic equation:

\[
\partial_t f(x, v, t) + \partial_x [vf] = Q[f],
\]

\[
Q[f] = U(x, v) \int \omega(v') f(x, v', t) \, dv' - \omega f
\]

\(\omega\): scattering frequency

Make the distribution \(U\) dependent on some macroscopic functional of the part density

\[
U(x, v) = U_S(x, v, s),
\]

\[
S = \begin{pmatrix}
\int f(x, v, t) \, dv \\
\int f(x, v, t) \, dxv \\
\int_1^{x} dx \int dv \, f(x, v, t)
\end{pmatrix}
\]
Instead of a clearing function (a flux or a group velocity) \( v(\rho) = \frac{F(\rho)}{\rho} \) we have a velocity distribution \( U_S(x, v, s) \, dv \), given in terms of a conditional probability.

\[
U_S(x, v, s) \, dv = dP[\eta = v | S(x, t) = s]
\]
Velocity distribution at stage 38 of the process (depending on the local WIP)
The update frequency $\omega(v)$

Mean free path:

$M \gg 1 \text{ stages} \Rightarrow \frac{v}{\omega} = \frac{1}{M} \Rightarrow \omega(v) = Mv, \quad M \gg 1$

The kinetic model:

$$\partial_t f(x, v, t) + \partial_x[vf] = \frac{1}{\varepsilon} Q[f],$$

$$Q[f] = U(x, v, S) \int \omega(v')f(x, v', t) \, dv' - \omega f, \quad \varepsilon = \frac{1}{M} \ll 1$$

$$S = \begin{pmatrix} \int f(x, v, t) \, dv & \text{or} \\ \int f(x, v, t) \, dxdv & \text{or} \\ \int_x^1 dx \int dv \, f(x, v, t) \end{pmatrix}$$
Chapman - Enskog for large time scales:

Split again the slow and fast time scales to obtain diffusion equation for the macroscopic density.

\[ f = f_0 + \varepsilon f_1, \quad f_0 = Pf, \quad \varepsilon f_1 = (I - P)f \]

\[ Pf = \frac{\rho U(x, v, S)}{\omega}, \quad \rho = \int f \, dv \]

\[ \partial_t \rho(x, t) + \partial_x [\langle V \rangle \rho - \varepsilon D \partial_x \rho] = 0, \quad D = \frac{\sigma(U)}{\langle U \rangle} \]

The amount of stochasticity in the system is given by the variation coefficient \( \frac{\sigma(U)}{\langle U \rangle} \).
Mean velocity as a function of total WIP
Diffusion coefficient as function of total WIP
Some Results 87

- Toy factory - Comparison (WIP) to discrete event simulations
- 26 processing steps, 200 machines, FIFO
Left: DES: (60,000 lots, 100 realizations), Right: CL
Part 1: Traffic flow vs. production systems

Part 2: First principle models.
  • Simple deterministic first principle models → hyperbolic conservation laws.
  • Stochasticity and kinetic models.
  • Mean field theory, long time averages and diffusive corrections.

Part 3: Clearing distributions from observed data.

Part 4:
  • The problem with diffusive corrections due to stochastic fluctuations. Hyperbolic relaxation models.
The basic problem:

▶ A diffusion equation (as a result of Chapman - Enskog) propagates information in both directions.

▶ Parts (or drivers in traffic flow) do not react to what is happening behind them.

▶ This is an artifact of the Chapman - Enskog procedure which transforms diffusion (arising from the random fluctuations in the flow) in velocity into spatial diffusion in a macroscopic limit.

▶ General problem for directional flows and fluctuations.
In practice, the equation

\[
\partial_t \rho(t, x) + \partial_x \{ac_{up}(x)[1 - \exp(-\frac{v_0\rho}{c_{up}})] - a\varepsilon D(\rho)\nabla^2 \partial_x \rho\}
\]

needs a boundary condition at \(x = 1\) and there is no 'physics' to determine this condition.
Solve the kinetic equation by a moment closure, taking additional (not conserved) moments.

⇒ a hyperbolic system, still containing $\varepsilon$.

Close the moment hierarchy by an ansatz, such that $\varepsilon \to 0$ asymptotics on the macroscopic level would reproduce the diffusion picture.

Don’t do it. Use the hyperbolic model instead.

Natalini, Jin, Slemrod ('95): Regularization of the Burnett and super-Burnett equations.
The Chapman - Enskog expansion revisited:

\[ \partial_t g + \partial_x (\tilde{V} g) = \frac{1}{\varepsilon} Q[g] \]

\[ g = g_0 + \varepsilon g_1 \quad g_0 = P g = \rho(x, t) G_0(x, C'), \quad \varepsilon g_1 = (I - P) g \]

(1) \[ \partial_t (\rho G_0) + \partial_x [P \tilde{V} (\rho G_0 + \varepsilon g_1)] = 0 \]

(2) \[ \varepsilon \partial_t g_1 + (I - P) \partial_x [\tilde{V} (\rho G_0 + \varepsilon g_1)] = Q[g_1] \]
In standard C.E. equation (2) is solved asymptotically, replacing it by

\[(2) \ (I - P)\partial_x[\tilde{V}(\rho G_0 + \varepsilon g_1)] = Q[g_1]\]

\[(1) \ \partial_t (\rho G_0) + \partial_x[P\tilde{V}(\rho G_0 + \varepsilon g_1)] = 0\]

which gives a diffusion term in the equation for \( \rho \)

\[\partial_t \rho + \partial_x[\rho\langle \tilde{V} G_0 \rangle + \varepsilon \langle g_1 \rangle] = 0\]
Alternative:

- Solve the original equation (2)

\[(2) \varepsilon \partial_t g_1 + (I - P) \partial_x [\tilde{V}(\rho G_0 + \varepsilon g_1)] = Q[g_1]\]

by a moment closure and close with an ansatz arising from the C.E. asymptotics.

- This gives a system which still contains the fast time scale \(O(\frac{1}{\varepsilon})\) terms.

- The system should be hyperbolic and (formally) reduce to the classical C.E. diffusion equation for \(\varepsilon \to 0\).
\[ \partial_t \rho + \partial_x (u \rho) = 0, \quad \partial_t (u \rho) + \partial_x [u^2 \rho + P \rho] = \frac{\rho}{\varepsilon} (u_0 - u) \]

\[ u \rho = \int \tilde{V} (\rho G_0 + \varepsilon g_1) \, dC, \quad u^2 \rho + P \rho = \int \tilde{V}^2 (\rho G_0 + \varepsilon g_1) \, dC \]

\( P \): pressure (from the closure).

\[ \text{Close with the asymptotic form given by Chapman - Enskog:} \]

\[ \rho u^2 + \rho P = \int \tilde{V}^2 (\rho G_0 + \varepsilon g_1)(x, C, t) \, dC \]

\( g_1 \) taken from the classical Chapman - Enskog procedure.

\[ \tilde{V} \rho G_0 = Q[g_1] \]
Asymptotics on the hyperbolic level:

\[ u(x, t) = u_0 - \frac{\varepsilon}{\rho} \partial_x[P\rho] \]

gives up to order \( \varepsilon^2 \) the same equation as the standard Chapman Enskog expansion.

The characteristic speeds of the second order system

\[
\partial_t \rho + \partial_x(u \rho) = 0, \quad \partial_t(u \rho) + \partial_x[u^2 \rho + P \rho] = \frac{\rho}{\varepsilon}(u_0 - u)
\]

are non-negative for \( \varepsilon << 1 \).
Remark

\[ \partial_t \rho + \partial_x (u \rho) = 0, \quad \partial_t (u \rho) + \partial_x [u^2 \rho + P \rho] = \frac{p}{\varepsilon} (u_0 - u) \]

- The numerical solution of the relaxation model causes no additional difficulties, since the (local) relaxation term can be discretized implicitly.

- Since the characteristics all point to the right, there is no boundary condition needed at \( x = L \).

- The boundary condition for the energy \( u^2 \rho + P \rho \) has to be computed from the equilibrium energy in the standard C.E. expansion.
The sign of the phase speeds has to be proven individually for a certain velocity profile \( \tilde{V} \) and a certain collision operator \( Q \).

In the nonlinear case \( \tilde{V} = \tilde{V}(\rho) \) we can guarantee the sign only in the limit \( \varepsilon \to 0 \).