The Derivation of Fluid Models from Point-Particle Dynamics

Brent Young
Rutgers University

bojy@math.rutgers.edu

August 28, 2010

Introduction: Classical Point-Particle Dynamics

In order to completely describe the state of a system of N identical particles moving under the influence of their own mutual interactions, we have to keep track of 6 numbers for each particle - 3 for its position, and 3 for its velocity (or momentum). Of course, we also need to know the type of interaction. We make the following convenient choices:

Introduction: Classical Point-Particle Dynamics

In order to completely describe the state of a system of N identical particles moving under the influence of their own mutual interactions, we have to keep track of 6 numbers for each particle - 3 for its position, and 3 for its velocity (or momentum). Of course, we also need to know the type of interaction. We make the following convenient choices:

- the mass of each particle is $\frac{1}{N}$ (so that the whole system has mass 1),

Introduction: Classical Point-Particle Dynamics

In order to completely describe the state of a system of N identical particles moving under the influence of their own mutual interactions, we have to keep track of 6 numbers for each particle - 3 for its position, and 3 for its velocity (or momentum). Of course, we also need to know the type of interaction. We make the following convenient choices:

- the mass of each particle is $\frac{1}{N}$ (so that the whole system has mass 1),
- we assume the particles interact via Newtonian gravitation,

Introduction: Classical Point-Particle Dynamics

In order to completely describe the state of a system of N identical particles moving under the influence of their own mutual interactions, we have to keep track of 6 numbers for each particle - 3 for its position, and 3 for its velocity (or momentum). Of course, we also need to know the type of interaction. We make the following convenient choices:

- the mass of each particle is $\frac{1}{N}$ (so that the whole system has mass 1),
- we assume the particles interact via Newtonian gravitation,
- we choose units so that certain physical constants are 1 (e.g. G),

Introduction: Classical Point-Particle Dynamics

In order to completely describe the state of a system of N identical particles moving under the influence of their own mutual interactions, we have to keep track of 6 numbers for each particle - 3 for its position, and 3 for its velocity (or momentum). Of course, we also need to know the type of interaction. We make the following convenient choices:

- the mass of each particle is $\frac{1}{N}$ (so that the whole system has mass 1),
- we assume the particles interact via Newtonian gravitation,
- we choose units so that certain physical constants are 1 (e.g. G),
- the position and velocity of the $\imath$-th particle will be denoted $q_\imath$ and $v_\imath$ respectively.

With these choices, the motion of our particles is governed by $6N$ coupled ordinary differential equations:

$$\begin{aligned}\dot{q}_\imath(t) &= v_\imath(t)\\ \dot{v}_\imath(t) &= \frac{1}{N}\sum_{\jmath\neq\imath}\frac{q_\jmath(t)-q_\imath(t)}{|q_\jmath(t)-q_\imath(t)|^3}\end{aligned}$$

with initial conditions $q_\imath(0)=\mathring{q}_\imath$ and $v_\imath(0)=\mathring{v}_\imath$ for each index $\imath$ (where the "$\cdot$" is standard shorthand for d/dt).

Introduction: Fluid Models

When using fluid models, we assume that the material under study is a continuum described by some sufficiently nice non-negative function that evolves in time. For such models, we need to establish a partial differential equation that describes the evolution of the system.

Introduction: Fluid Models

When using fluid models, we assume that the material under study is a continuum described by some sufficiently nice non-negative function that evolves in time. For such models, we need to establish a partial differential equation that describes the evolution of the system.

- Since each particle in our system has two degrees of freedom at all times, we look for a time-dependent function on $\mathbb{R}^6$ - denoted $f_t(v, q)$.

Introduction: Fluid Models

When using fluid models, we assume that the material under study is a continuum described by some sufficiently nice non-negative function that evolves in time. For such models, we need to establish a partial differential equation that describes the evolution of the system.

- Since each particle in our system has two degrees of freedom at all times, we look for a time-dependent function on $\mathbb{R}^6$ - denoted $f_t(v, q)$.
- Intuitively, $f_t(v, q)$ represents the amount of fluid in a tiny volume centered at position q with velocity v.

Introduction: Fluid Models

When using fluid models, we assume that the material under study is a continuum described by some sufficiently nice non-negative function that evolves in time. For such models, we need to establish a partial differential equation that describes the evolution of the system.

- Since each particle in our system has two degrees of freedom at all times, we look for a time-dependent function on $\mathbb{R}^6$ - denoted $f_t(v, q)$.
- Intuitively, $f_t(v, q)$ represents the amount of fluid in a tiny volume centered at position q with velocity v.
- As we want the total mass to be one, we require $\iint f_t(v, q)\, d^3v\, d^3q = 1$.

Introduction: Fluid Models

When using fluid models, we assume that the material under study is a continuum described by some sufficiently nice non-negative function that evolves in time. For such models, we need to establish a partial differential equation that describes the evolution of the system.

- Since each particle in our system has two degrees of freedom at all times, we look for a time-dependent function on $\mathbb{R}^6$ - denoted $f_t(v, q)$.
- Intuitively, $f_t(v, q)$ represents the amount of fluid in a tiny volume centered at position q with velocity v.
- As we want the total mass to be one, we require $\iint f_t(v, q)\, d^3v\, d^3q = 1$.
- When taking derivatives of f_t, we write ∇_q to indicate the 3-dimensional gradient w.r.t. the position variables (and similarly for the velocity variables).

The difficult question is which fluid model to choose? In other words, what PDE for f_t should we expect closely approximates our given point-particle dynamics? It turns out that the correct one is a coupled system of PDEs known as the Vlasov-Poisson system:

$$\partial_t f_t(v,q) + v \cdot \nabla_q f_t(v,q) - \nabla_q \phi_t(q) \cdot \nabla_v f_t(v,q) = 0,$$
$$\triangle_q \phi_t(q) = 4\pi \int f_t(v,q)\, d^3v,$$
$$\phi_t(q) \asymp -\frac{1}{|q|} \text{ as } |q| \to \infty$$

with $f_0(v,q) = \mathring{f}(v,q)$.

The naming is a little contentious as Jeans seems to have been the first to write down this system (in 1915). Though Vlasov's name is commonly used, his results were published (independently it seems) in 1938.

Our intuition is that a fluid model should be a good approximation to the underlying particle motion when the number of particles is sufficiently large and when our scale of observation is large compared to the typical particle separation. In order to make this intuition into solid mathematics, we have to overcome three hurdles:

Our intuition is that a fluid model should be a good approximation to the underlying particle motion when the number of particles is sufficiently large and when our scale of observation is large compared to the typical particle separation. In order to make this intuition into solid mathematics, we have to overcome three hurdles:

1) In order to say that a collection of particles (in $\mathbb{R}^6$) is somehow close to a continuous function (over $\mathbb{R}^6$), we must first realize these two very different kinds of objects as elements in the same space,

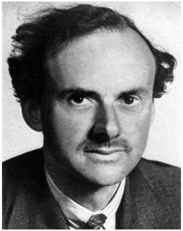
Our intuition is that a fluid model should be a good approximation to the underlying particle motion when the number of particles is sufficiently large and when our scale of observation is large compared to the typical particle separation. In order to make this intuition into solid mathematics, we have to overcome three hurdles:

1) In order to say that a collection of particles (in $\mathbb{R}^6$) is somehow close to a continuous function (over $\mathbb{R}^6$), we must first realize these two very different kinds of objects as elements in the same space,
2) Once we find such a generalized class of objects, we need some measure of how close two elements in this space are to one another (i.e. a metric),

Our intuition is that a fluid model should be a good approximation to the underlying particle motion when the number of particles is sufficiently large and when our scale of observation is large compared to the typical particle separation. In order to make this intuition into solid mathematics, we have to overcome three hurdles:

1) In order to say that a collection of particles (in $\mathbb{R}^6$) is somehow close to a continuous function (over $\mathbb{R}^6$), we must first realize these two very different kinds of objects as elements in the same space,
2) Once we find such a generalized class of objects, we need some measure of how close two elements in this space are to one another (i.e. a metric),
3) After laying the foundations above, we have to do the actual analysis proving that an evolving function closely approximates a collection of moving points.

Schwartz Distributions



P.A.M. Dirac was a British theoretical physicist in the first half of the Twentieth Century. His most famous contributions to science lie in the field of Quantum Physics. Of the many contributions he made to mathematical physics, the most well known is the Dirac Delta "Function." Dirac envisioned a function $\delta(x)$ with the following properties:

- $\int \delta(x)\, dx = 1$
- For a sufficiently nice function $f(x)$ (say continuous in some neighborhood of the origin),

$$\int f(x)\delta(x)\, dx = f(0).$$

- Dirac clearly indicated that in some sense we can think of the Delta as the derivative of the Heaviside Function ($H(x) = 1$ for $x \geq 0$ and 0 otherwise).

- Dirac clearly indicated that in some sense we can think of the Delta as the derivative of the Heaviside Function ($H(x) = 1$ for $x \geq 0$ and 0 otherwise).
- This is strongly suggested by analogy to Riemann-Stieltjes integration.

- Dirac clearly indicated that in some sense we can think of the Delta as the derivative of the Heaviside Function ($H(x) = 1$ for $x \geq 0$ and 0 otherwise).
- This is strongly suggested by analogy to Riemann-Stieltjes integration.
- Unfortunately, mathematicians quickly pointed out that no reasonable function can have the properties Dirac listed.

- Dirac clearly indicated that in some sense we can think of the Delta as the derivative of the Heaviside Function ($H(x) = 1$ for $x \geq 0$ and 0 otherwise).
- This is strongly suggested by analogy to Riemann-Stieltjes integration.
- Unfortunately, mathematicians quickly pointed out that no reasonable function can have the properties Dirac listed.
- Whatever the Dirac Delta is, it seems a sensible object. The problem is coming to terms with it in a mathematically acceptable fashion.

- Dirac clearly indicated that in some sense we can think of the Delta as the derivative of the Heaviside Function ($H(x) = 1$ for $x \geq 0$ and 0 otherwise).
- This is strongly suggested by analogy to Riemann-Stieltjes integration.
- Unfortunately, mathematicians quickly pointed out that no reasonable function can have the properties Dirac listed.
- Whatever the Dirac Delta is, it seems a sensible object. The problem is coming to terms with it in a mathematically acceptable fashion.
- As the Dirac Delta is not a function, we will write $< \delta, f >= f(0)$.

- Dirac clearly indicated that in some sense we can think of the Delta as the derivative of the Heaviside Function ($H(x) = 1$ for $x \geq 0$ and 0 otherwise).
- This is strongly suggested by analogy to Riemann-Stieltjes integration.
- Unfortunately, mathematicians quickly pointed out that no reasonable function can have the properties Dirac listed.
- Whatever the Dirac Delta is, it seems a sensible object. The problem is coming to terms with it in a mathematically acceptable fashion.
- As the Dirac Delta is not a function, we will write $< \delta, f >= f(0)$.
- The solution came from Laurent Schwartz in 1945 (for which he received the Fields Medal in 1950).

Though Schwartz was not directly working to understand the Dirac Delta, his work deals with objects that have the following basic properties:

Though Schwartz was not directly working to understand the Dirac Delta, his work deals with objects that have the following basic properties:

- the objects in question take in *sufficiently nice* functions and return numbers (such objects are typically called functionals),

Though Schwartz was not directly working to understand the Dirac Delta, his work deals with objects that have the following basic properties:

- the objects in question take in *sufficiently nice* functions and return numbers (such objects are typically called functionals),
- the objects in question are linear: for functions f, g and scalars α, β

$$<\delta, \alpha f + \beta g> = \alpha < \delta, f > + \beta < \delta, g >,$$

Though Schwartz was not directly working to understand the Dirac Delta, his work deals with objects that have the following basic properties:

- the objects in question take in *sufficiently nice* functions and return numbers (such objects are typically called functionals),
- the objects in question are linear: for functions f, g and scalars α, β

$$< \delta, \alpha f + \beta g >= \alpha < \delta, f > + \beta < \delta, g >,$$

- the size of the output can be controlled by the size of the input (so the objects are said to be bounded).

Though Schwartz was not directly working to understand the Dirac Delta, his work deals with objects that have the following basic properties:

- the objects in question take in *sufficiently nice* functions and return numbers (such objects are typically called functionals),
- the objects in question are linear: for functions f, g and scalars α, β
$$< \delta, \alpha f + \beta g >= \alpha < \delta, f > + \beta < \delta, g >,$$
- the size of the output can be controlled by the size of the input (so the objects are said to be bounded).

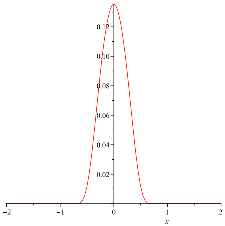
These objects are referred to as *bounded linear functionals* on whatever space of functions you choose.
Schwartz realized that by restricting to a very nice class of functions not only do you get a correspondingly wider class of functionals, but many of the properties enjoyed by the functions can be transferred over to the functionals.

Our first task is to choose an appropriate space of functions. We will restrict our attention to functions over $\mathbb{R}$ (the generalization to $\mathbb{R}^n$ is simple). Given our comments above we choose a very nice class of functions indeed:

$$\mathcal{C}_c^\infty(\mathbb{R}) = \{\psi : \mathbb{R} \to \mathbb{R} | \psi \in \mathcal{C}^\infty(\mathbb{R}) \ \& \ \psi \equiv 0 \text{ outside some compact set}\}.$$

This set of functions is referred to as the set of $\mathcal{C}^\infty$ *functions with compact support* on $\mathbb{R}$.

A prototypical example is provided by



$$\psi(x) = \begin{cases} e^{-(x+1)^{-2}} e^{-(x-1)^{-2}} & -1 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

In order to discuss *bounded* functionals on this space, we need to endow our class of functions with a topology - called the *Strict Inductive Limit Topology*. Instead of giving the (gory) details of open sets and the like, it suffices to say what we mean for a sequence $\{\psi_n\}_{n=1}^{\infty}$ to converge to ψ in this topology.
$\psi_n \to \psi$ in the Strict Inductive Limit Topology iff

1) there exists a compact set C so that each ψ_n and ψ are identically zero outside C.
2) for every m, $\lim_{n\to\infty} \sup_{\mathbb{R}} |\psi_n^{(m)}(x) - \psi^{(m)}(x)| = 0$ (uniform convergence in all derivatives).

We denote by $\mathcal{D}(\mathbb{R})$ the set $\mathcal{C}_c^{\infty}(\mathbb{R})$ equipped with this topology. This is referred to as the space of *Test Functions*.

$\mathcal{D}'(\mathbb{R})$ denotes the collection of bounded linear functionals on the test function space $\mathcal{D}(\mathbb{R})$. These objects are called *Schwartz Distributions* (or just distributions if there is no chance for confusion). We present some basic examples. It is standard to write $< f, \psi >$ for the action of the functional f on the test function ψ (later we will change this notation to something more suggestive).

$\mathcal{D}'(\mathbb{R})$ denotes the collection of bounded linear functionals on the test function space $\mathcal{D}(\mathbb{R})$. These objects are called *Schwartz Distributions* (or just distributions if there is no chance for confusion). We present some basic examples. It is standard to write $< f, \psi >$ for the action of the functional f on the test function ψ (later we will change this notation to something more suggestive).

- The Dirac Delta is in $\mathcal{D}'(\mathbb{R})$: $< \delta, \psi >= \psi(0)$. This is clearly linear. Boundedness follows as

$\psi(0) \leq \sup_{\mathbb{R}} |\psi(x)|$.

$\mathcal{D}'(\mathbb{R})$ denotes the collection of bounded linear functionals on the test function space $\mathcal{D}(\mathbb{R})$. These objects are called *Schwartz Distributions* (or just distributions if there is no chance for confusion). We present some basic examples. It is standard to write $< f, \psi >$ for the action of the functional f on the test function ψ (later we will change this notation to something more suggestive).

- The Dirac Delta is in $\mathcal{D}'(\mathbb{R})$: $< \delta, \psi >= \psi(0)$. This is clearly linear. Boundedness follows as $\psi(0) \leq \sup_{\mathbb{R}} |\psi(x)|$.
- Let f be any continuous function. Then we define the distribution I_f by

$$< I_f, \psi >= \int f(x)\psi(x)\, dx.$$

  This is clearly linear. Boundedness is also easy to show.

- Building on our last example, let f be any function such that for any $-\infty < a < b < \infty$ we have

$$\int_a^b |f(x)|\, dx < \infty.$$

Then we can define I_f as in the last example. Such functions are said to be *locally integrable*.

- Building on our last example, let f be any function such that for any $-\infty < a < b < \infty$ we have

$$\int_a^b |f(x)| \, dx < \infty.$$

Then we can define I_f as in the last example. Such functions are said to be *locally integrable*.

- Often, the distributions arising from integration against locally integrable functions are called *regular distributions*.

- Building on our last example, let f be any function such that for any $-\infty < a < b < \infty$ we have

$$\int_a^b |f(x)|\, dx < \infty.$$

Then we can define I_f as in the last example. Such functions are said to be *locally integrable*.

- Often, the distributions arising from integration against locally integrable functions are called *regular distributions*.
- Define the distribution δ' by $<\delta', \psi> = -\psi'(0)$. This is clearly linear. Boundedness follows exactly like the Dirac Delta, except here the uniform bound is on the derivative of ψ.

Making rigorous sense of the Dirac Delta and unifying it with the seemingly much different concept of function is just the beginning. The amazing insight of Schwartz was that we can extend many operations defined for the test functions to analogous properties for the distributions. The usual recipe for accomplishing this is given below.

Making rigorous sense of the Dirac Delta and unifying it with the seemingly much different concept of function is just the beginning. The amazing insight of Schwartz was that we can extend many operations defined for the test functions to analogous properties for the distributions. The usual recipe for accomplishing this is given below.

- Start with distributions for which the operation in question is already well-defined. Typically this involves using the regular distributions. Remember that EVERY $\mathcal{C}^\infty$ function is a distribution!

Making rigorous sense of the Dirac Delta and unifying it with the seemingly much different concept of function is just the beginning. The amazing insight of Schwartz was that we can extend many operations defined for the test functions to analogous properties for the distributions. The usual recipe for accomplishing this is given below.

- Start with distributions for which the operation in question is already well-defined. Typically this involves using the regular distributions. Remember that EVERY $\mathcal{C}^{\infty}$ function is a distribution!
- Manipulate the resulting integral so as to pass the operation over to the test function. Typically this involves integration by parts, changing variables, changing the order of integration, etc.

Making rigorous sense of the Dirac Delta and unifying it with the seemingly much different concept of function is just the beginning. The amazing insight of Schwartz was that we can extend many operations defined for the test functions to analogous properties for the distributions. The usual recipe for accomplishing this is given below.

- Start with distributions for which the operation in question is already well-defined. Typically this involves using the regular distributions. Remember that EVERY $\mathcal{C}^\infty$ function is a distribution!
- Manipulate the resulting integral so as to pass the operation over to the test function. Typically this involves integration by parts, changing variables, changing the order of integration, etc.
- Use the result to define the operation for all distributions.

Making rigorous sense of the Dirac Delta and unifying it with the seemingly much different concept of function is just the beginning. The amazing insight of Schwartz was that we can extend many operations defined for the test functions to analogous properties for the distributions. The usual recipe for accomplishing this is given below.

- Start with distributions for which the operation in question is already well-defined. Typically this involves using the regular distributions. Remember that EVERY $\mathcal{C}^{\infty}$ function is a distribution!
- Manipulate the resulting integral so as to pass the operation over to the test function. Typically this involves integration by parts, changing variables, changing the order of integration, etc.
- Use the result to define the operation for all distributions.

We illustrate the process by showing how to define derivatives of distributions.

Derivatives of Distributions

We start with a function f that already has a derivative in the usual sense. To ensure that our new notion of differentiation extends the old notion, we **MUST** have $I_f' = I_{f'}$ (we drop the $I_.$ notion from now on).

Derivatives of Distributions

We start with a function f that already has a derivative in the usual sense. To ensure that our new notion of differentiation extends the old notion, we **MUST** have $I_f' = I_{f'}$ (we drop the $I_.$ notion from now on).

- $$\begin{aligned} < f', \psi > &= \int f'(x)\psi(x)\,dx \\ &= f(x)\psi(x)|_{-\infty}^{\infty} - \int f(x)\psi'(x)\,dx \\ &= -\int f(x)\psi'(x)\,dx \\ &= - < f, \psi' > \end{aligned}$$

Derivatives of Distributions

We start with a function f that already has a derivative in the usual sense. To ensure that our new notion of differentiation extends the old notion, we **MUST** have $I_f' = I_{f'}$ (we drop the $I_.$ notion from now on).

- $$\begin{aligned} < f', \psi > &= \int f'(x)\psi(x)\, dx \\ &= f(x)\psi(x)|_{-\infty}^{\infty} - \int f(x)\psi'(x)\, dx \\ &= -\int f(x)\psi'(x)\, dx \\ &= -< f, \psi' > \end{aligned}$$

- So we now **DEFINE** for any distribution f

$$< f', \psi > \equiv - < f, \psi' > .$$

- Repeated integration by parts gives the natural definition of the *n*-th derivative of a distribution

$$< f^{(n)}, \psi > \equiv (-1)^n < f, \psi^{(n)} > .$$

- Repeated integration by parts gives the natural definition of the *n*-th derivative of a distribution

$$< f^{(n)}, \psi > \equiv (-1)^n < f, \psi^{(n)} > .$$

- Note that **EVERY DISTRIBUTION IS INFINITELY DIFFERENTIABLE** in this new sense!

We illustrate the process with a classic example, $f(x) = |x|$.

We illustrate the process with a classic example, $f(x) = |x|$.

- $< f', \psi >= - < f, \psi' >$

We illustrate the process with a classic example, $f(x) = |x|$.

- $< f', \psi > = - < f, \psi' >$
-

$$\begin{aligned} - < f, \psi' > &= - \int |x| \psi'(x)\, dx \\ &= \int_{-\infty}^{0} x\psi'(x)\, dx - \int_{0}^{\infty} x\psi'(x)\, dx \\ &= - \int_{-\infty}^{0} \psi(x)\, dx + \int_{0}^{\infty} \psi(x)\, dx \\ &= \int h(x)\psi(x)\, dx = < h, \psi > \end{aligned}$$

We illustrate the process with a classic example, $f(x) = |x|$.

- $< f', \psi >= - < f, \psi' >$
-

$$
\begin{aligned}
- < f, \psi' > &= - \int |x| \psi'(x)\, dx \\
&= \int_{-\infty}^{0} x\psi'(x)\, dx - \int_{0}^{\infty} x\psi'(x)\, dx \\
&= - \int_{-\infty}^{0} \psi(x)\, dx + \int_{0}^{\infty} \psi(x)\, dx \\
&= \int h(x)\psi(x)\, dx =< h, \psi >
\end{aligned}
$$

- Here, $h(x) = \begin{cases} -1 & x < 0 \\ 1 & x \geq 0 \end{cases}$

We illustrate the process with a classic example, $f(x) = |x|$.

- $< f', \psi > = - < f, \psi' >$
-

$$
\begin{aligned}
- < f, \psi' > &= - \int |x| \psi'(x)\, dx \\
&= \int_{-\infty}^{0} x\psi'(x)\, dx - \int_{0}^{\infty} x\psi'(x)\, dx \\
&= - \int_{-\infty}^{0} \psi(x)\, dx + \int_{0}^{\infty} \psi(x)\, dx \\
&= \int h(x)\psi(x)\, dx = < h, \psi >
\end{aligned}
$$

- Here, $h(x) = \begin{cases} -1 & x < 0 \\ 1 & x \geq 0 \end{cases}$

So $f'(x) = h(x)$. Note that the value of h at the origin is completely irrelevant!

What about the derivative of $h(x)$?

- $< h', \psi >= - < h, \psi' >$

What about the derivative of $h(x)$?

- $< h', \psi > = - < h, \psi' >$
-

$$\begin{aligned} - < h, \psi' > &= - \int h(x)\psi'(x)\, dx \\ &= \int_{-\infty}^{0} \psi'(x)\, dx - \int_{0}^{\infty} \psi'(x)\, dx \\ &= \psi(x)|_{-\infty}^{0} - \psi(x)|_{0}^{\infty} \\ &= 2\psi(0) \\ &= < 2\delta, \psi > \end{aligned}$$

What about the derivative of $h(x)$?

- $< h', \psi >= - < h, \psi' >$
-

$$
\begin{aligned}
- < h, \psi' > &= - \int h(x)\psi'(x)\, dx \\
&= \int_{-\infty}^{0} \psi'(x)\, dx - \int_{0}^{\infty} \psi'(x)\, dx \\
&= \psi(x)|_{-\infty}^{0} - \psi(x)|_{0}^{\infty} \\
&= 2\psi(0) \\
&= < 2\delta, \psi >
\end{aligned}
$$

So $h'(x) = 2\delta(x)$! The coefficient 2 corresponds to the jump in h (from -1 to 1) being 2 units large.

Finally, we can give a definitive answer to the first of our problems: we characterize both the point particle description of a state and the fluid description as Schwartz distributions.

Finally, we can give a definitive answer to the first of our problems: we characterize both the point particle description of a state and the fluid description as Schwartz distributions.

- Specifically, the fluid model is taken to be integration against the function f_t (of course, now our test functions are over $\mathbb{R}^6$).

Finally, we can give a definitive answer to the first of our problems: we characterize both the point particle description of a state and the fluid description as Schwartz distributions.

- Specifically, the fluid model is taken to be integration against the function f_t (of course, now our test functions are over $\mathbb{R}^6$).
- The particles are modeled by a sum of Dirac Deltas:

$$\Delta_t^N(v,q) = \frac{1}{N}\sum_{i=1}^{N} \delta(v - v_i(t))\delta(q - q_i(t)).$$

Finally, we can give a definitive answer to the first of our problems: we characterize both the point particle description of a state and the fluid description as Schwartz distributions.

- Specifically, the fluid model is taken to be integration against the function f_t (of course, now our test functions are over $\mathbb{R}^6$).
- The particles are modeled by a sum of Dirac Deltas:

$$\Delta_t^N(v,q) = \frac{1}{N}\sum_{i=1}^{N}\delta(v - v_i(t))\delta(q - q_i(t)).$$

- The action of one of these Delta terms on a test function (now in $\mathbb{R}^6$) is

$$< \delta(v - v_\imath(t))\delta(q - q_\imath(t)), \psi >= \psi(v_\imath(t), q_\imath(t)).$$

Finally, we can give a definitive answer to the first of our problems: we characterize both the point particle description of a state and the fluid description as Schwartz distributions.

- Specifically, the fluid model is taken to be integration against the function f_t (of course, now our test functions are over $\mathbb{R}^6$).
- The particles are modeled by a sum of Dirac Deltas:

$$\Delta_t^N(v,q) = \frac{1}{N}\sum_{i=1}^{N}\delta(v - v_i(t))\delta(q - q_i(t)).$$

- The action of one of these Delta terms on a test function (now in $\mathbb{R}^6$) is

$$< \delta(v - v_\imath(t))\delta(q - q_\imath(t)), \psi >= \psi(v_\imath(t), q_\imath(t)).$$

- Moreover, by making the total mass of the system 1, both of these objects can also be thought of as *probability measures* on $\mathbb{R}^6$.

Finally, we can give a definitive answer to the first of our problems: we characterize both the point particle description of a state and the fluid description as Schwartz distributions.

- Specifically, the fluid model is taken to be integration against the function f_t (of course, now our test functions are over $\mathbb{R}^6$).
- The particles are modeled by a sum of Dirac Deltas:

$$\Delta_t^N(v,q) = \frac{1}{N}\sum_{i=1}^{N}\delta(v - v_i(t))\delta(q - q_i(t)).$$

- The action of one of these Delta terms on a test function (now in $\mathbb{R}^6$) is

$$< \delta(v - v_\imath(t))\delta(q - q_\imath(t)), \psi >= \psi(v_\imath(t), q_\imath(t)).$$

- Moreover, by making the total mass of the system 1, both of these objects can also be thought of as *probability measures* on $\mathbb{R}^6$.
- It is important to note that even though we characterize the system as a probability measure, *there is absolutely no randomness* once we specify an initial condition!

Now, using our generalized version of differentiation for distributions, we can find a PDE for the point particle distributions. With a little effort, we find :

$$\partial_t \Delta_t^N + v \cdot \nabla_q \Delta_t^N - \nabla_q \phi_t \cdot \nabla_v \Delta_t^N = 0,$$
$$\triangle_q \phi_t = 4\pi \int \Delta_t^N d^3 v$$

which is the same PDE as we have for f_t!

The Dual Bounded Lipschitz Metric

Now that we have established what sort of objects we should use to model our system, we turn to the question of convergence. How should we make sense of the notion that $\Delta_t^N \to f_t$ as $N \to \infty$? There is a natural topology on $\mathcal{D}'$ - called the *weak-$\star$ topology*:

$$\lim_{N\to\infty} \Delta_t^N = f_t \quad \text{iff} \quad \lim_{N\to\infty} < \Delta_t^N, \psi > = < f_t, \psi >$$

for every test function $\psi \in \mathcal{D}$.
This topology has two major problems:

The Dual Bounded Lipschitz Metric

Now that we have established what sort of objects we should use to model our system, we turn to the question of convergence. How should we make sense of the notion that $\Delta_t^N \to f_t$ as $N \to \infty$? There is a natural topology on $\mathcal{D}'$ - called the *weak-$\star$ topology*:

$$\lim_{N\to\infty} \Delta_t^N = f_t \quad \text{iff} \quad \lim_{N\to\infty} < \Delta_t^N, \psi > = < f_t, \psi >$$

for every test function $\psi \in \mathcal{D}$.
This topology has two major problems:

- 1) It is NOT metrizable,

The Dual Bounded Lipschitz Metric

Now that we have established what sort of objects we should use to model our system, we turn to the question of convergence. How should we make sense of the notion that $\Delta_t^N \to f_t$ as $N \to \infty$? There is a natural topology on $\mathcal{D}'$ - called the *weak-$\star$ topology*:

$$\lim_{N\to\infty} \Delta_t^N = f_t \quad \text{iff} \quad \lim_{N\to\infty} < \Delta_t^N, \psi > = < f_t, \psi >$$

for every test function $\psi \in \mathcal{D}$.

This topology has two major problems:

- 1) It is NOT metrizable,
- 2) Convergence in this topology is likely to be VERY slow as a test function can be narrowly supported.

The topology we use makes use of the fact that we are dealing exclusively with probability measures. The intuitive idea is to limit ourselves to functions that cannot be so narrowly concentrated. To do so, we have to use some criterion that will ensure that the first derivatives of test functions cannot be too large.

The topology we use makes use of the fact that we are dealing exclusively with probability measures. The intuitive idea is to limit ourselves to functions that cannot be so narrowly concentrated. To do so, we have to use some criterion that will ensure that the first derivatives of test functions cannot be too large.

- We will actually use functions that are merely continuous (and bounded). There is a potential issue since we defined our distributions with a much more restrictive class of test functions. This is no problem for us as the Dirac Delta makes sense for any function continuous around the origin and f_t integrates to one for all t.

The topology we use makes use of the fact that we are dealing exclusively with probability measures. The intuitive idea is to limit ourselves to functions that cannot be so narrowly concentrated. To do so, we have to use some criterion that will ensure that the first derivatives of test functions cannot be too large.

- We will actually use functions that are merely continuous (and bounded). There is a potential issue since we defined our distributions with a much more restrictive class of test functions. This is no problem for us as the Dirac Delta makes sense for any function continuous around the origin and f_t integrates to one for all t.
- Though the functions we consider will not necessarily have derivatives in the classical sense, we can still restrict the rapidity with which they can grow. The necessary tool for doing this is the Lipschitz constant.

Lipschitz Continuity

A function is said to be *Lipschitz Continuous* if there exists a constant $L > 0$ so that

$$|f(x) - f(y)| \quad \leq \quad L|x - y|$$

for all x and y in the domain of f. There is a corresponding semi-norm

$$[f]_\ell \quad = \quad \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}.$$

Clearly, f is Lipschitz Continuous iff $[f]_\ell < \infty$.

- Note that if f is differentiable at x and Lipschitz Continuous, then

$$|f'(x)| \leq [f]_\ell.$$

  Thus, the Lipschitz Constant $[f]_\ell$ effectively controls the rate at which the function can change!

- Note that if f is differentiable at x and Lipschitz Continuous, then

$$|f'(x)| \leq [f]_\ell.$$

  Thus, the Lipschitz Constant $[f]_\ell$ effectively controls the rate at which the function can change!

- We define a new class of functions

$$D(\Omega) = \{\varphi : \Omega \to [0,1] \mid [\varphi]_\ell \leq 1\}.$$

- Note that if f is differentiable at x and Lipschitz Continuous, then

$$|f'(x)| \leq [f]_\ell.$$

  Thus, the Lipschitz Constant $[f]_\ell$ effectively controls the rate at which the function can change!
- We define a new class of functions

$$D(\Omega) = \{\varphi : \Omega \to [0,1] \mid [\varphi]_\ell \leq 1\}.$$

- These functions are positive, bounded (by 1), and Lipschitz Continuous. The choice to bound these functions by 1 (both in absolute size and Lipschitz Constant) is arbitrary. As long as these quantities are bounded by some size, then we will get the same results in our application.

We define a metric on probability measures by

$$d_{bL^*}(\mu, \nu) = \sup_{\varphi \in D} \left| \int \varphi d(\mu - \nu) \right|.$$

This metric is called the *dual bounded Lipschitz metric*.

We define a metric on probability measures by

$$d_{bL^*}(\mu,\nu) = \sup_{\varphi \in D} \left| \int \varphi d(\mu - \nu) \right|.$$

This metric is called the *dual bounded Lipschitz metric*.

- Verifying that this is a metric requires a little work (but we will take this for granted without proof).

We define a metric on probability measures by

$$d_{bL^*}(\mu, \nu) = \sup_{\varphi \in D} \left| \int \varphi d(\mu - \nu) \right| .$$

This metric is called the *dual bounded Lipschitz metric*.

- Verifying that this is a metric requires a little work (but we will take this for granted without proof).
- The adjective "Lipschitz" in the name is obvious. The use of "bounded" is also clear since $\varphi \geq 0$ and the fact that we are working with probability measures gives

  $$0 \leq d_{bL^*}(\mu, \nu) \leq 1$$

  for any two probability measures μ and ν.

We define a metric on probability measures by

$$d_{bL^*}(\mu, \nu) = \sup_{\varphi \in D} \left| \int \varphi d(\mu - \nu) \right| .$$

This metric is called the *dual bounded Lipschitz metric*.

- Verifying that this is a metric requires a little work (but we will take this for granted without proof).
- The adjective “Lipschitz” in the name is obvious. The use of “bounded” is also clear since $\varphi \geq 0$ and the fact that we are working with probability measures gives
$$0 \leq d_{bL^*}(\mu, \nu) \leq 1$$
for any two probability measures μ and ν.
- “Dual” refers to the fact that this metric is dual (in a technical sense) to another metric (the so-called Wasserstein metric). This other metric is more difficult to work with in our particular application which is why we prefer the dual bounded Lipschitz metric.

Now that we have settled the first two issues mentioned at the outset, we can begin the actual analysis! To that end, we can finally state the main theorem:

Theorem (Neunzert 1975)

Let $\{\Delta_0^N\}_{N=1}^{\infty}$ be a sequence of discrete initial data, and let f_0 be a sufficiently nice function such that

$$\lim_{N\to\infty} d_{bL^*}(\Delta_0^N, f_0) = 0.$$

For all times $t > 0$, let $\{\Delta_t^N\}_{N=1}^{\infty}$ be the corresponding points evolved according to the Newtonian Gravitational dynamics, and let f_t be obtained from f_0 according to the Vlasov-Poisson PDE. Then we have

$$\lim_{N\to\infty} d_{bL^*}(\Delta_t^N, f_t) = 0.$$

Proving this requires three preliminary steps:

Now that we have settled the first two issues mentioned at the outset, we can begin the actual analysis! To that end, we can finally state the main theorem:

Theorem (Neunzert 1975)

Let $\{\Delta_0^N\}_{N=1}^{\infty}$ be a sequence of discrete initial data, and let f_0 be a sufficiently nice function such that

$$\lim_{N\to\infty} d_{bL^*}(\Delta_0^N, f_0) = 0.$$

For all times $t > 0$, let $\{\Delta_t^N\}_{N=1}^{\infty}$ be the corresponding points evolved according to the Newtonian Gravitational dynamics, and let f_t be obtained from f_0 according to the Vlasov-Poisson PDE. Then we have

$$\lim_{N\to\infty} d_{bL^*}(\Delta_t^N, f_t) = 0.$$

Proving this requires three preliminary steps:

- 1) Gronwall's Inequality,

Now that we have settled the first two issues mentioned at the outset, we can begin the actual analysis! To that end, we can finally state the main theorem:

Theorem (Neunzert 1975)

Let $\{\Delta_0^N\}_{N=1}^{\infty}$ be a sequence of discrete initial data, and let f_0 be a sufficiently nice function such that

$$\lim_{N\to\infty} d_{bL^*}(\Delta_0^N, f_0) = 0.$$

For all times $t > 0$, let $\{\Delta_t^N\}_{N=1}^{\infty}$ be the corresponding points evolved according to the Newtonian Gravitational dynamics, and let f_t be obtained from f_0 according to the Vlasov-Poisson PDE. Then we have

$$\lim_{N\to\infty} d_{bL^*}(\Delta_t^N, f_t) = 0.$$

Proving this requires three preliminary steps:

- 1) Gronwall's Inequality,
- 2) Regularizing the gravitational potential,

Now that we have settled the first two issues mentioned at the outset, we can begin the actual analysis! To that end, we can finally state the main theorem:

Theorem (Neunzert 1975)

Let $\{\Delta_0^N\}_{N=1}^{\infty}$ *be a sequence of discrete initial data, and let* f_0 *be a sufficiently nice function such that*

$$\lim_{N\to\infty} d_{bL^*}(\Delta_0^N, f_0) = 0.$$

For all times $t > 0$*, let* $\{\Delta_t^N\}_{N=1}^{\infty}$ *be the corresponding points evolved according to the Newtonian Gravitational dynamics, and let* f_t *be obtained from* f_0 *according to the Vlasov-Poisson PDE. Then we have*

$$\lim_{N\to\infty} d_{bL^*}(\Delta_t^N, f_t) = 0.$$

Proving this requires three preliminary steps:

- 1) Gronwall's Inequality,
- 2) Regularizing the gravitational potential,
- 3) The Fixed Point Characterization of our dynamics.

Theorem (Gronwall's Inequality)

Let α, β and u be real-valued functions defined on an interval $[0, T)$ where $\beta \geq 0$ and u are continuous and the negative part of α is integrable (on $[0, T)$). Then if for each $t \in [0, T)$

$$u(t) \leq \alpha(t) + \int_0^t \beta(\tau)u(\tau)d\tau$$

then we also have

$$u(t) \leq \alpha(t) + \int_0^t \alpha(\tau)\beta(\tau)\exp\left(\int_\tau^t \beta(s)ds\right)d\tau.$$

For a proof, consult most any reasonably advanced text on Differential Equations (or Wikipedia!).

For our purposes,

$$\alpha(t) = \int_0^t f(\tau)d\tau$$

(with $f \geq 0$) and $\beta(t) = L$. In this special case, we get the following corollary of Gronwall's Inequality (after an integration by parts and some simplification): If for each $t \in [0, T)$

$$u(t) \leq \int_0^t f(\tau)d\tau + L\int_0^t u(\tau)d\tau$$

then we have

$$u(t) \leq e^{Lt}\int_0^t f(\tau)e^{-L\tau}d\tau.$$

Regularization

The Newtonian gravitational force is singular at each of the point particles. This leads to serious problems in the derivation. To avoid this issue, we regularize the force to remove the singularities. There are many ways to do this, but the simplest is to replace the usual gravitational force:

$$G(x, y) = \frac{x - y}{|x - y|^3}$$

by the modified force

$$G_\epsilon(x, y) = \frac{x - y}{(|x - y|^2 + \epsilon)^{3/2}}$$

for some small $\epsilon > 0$. There is a result by Pfaffelmoser that the regularization can be removed at the end of the proof (i.e. we can allow $\epsilon \to 0$). From now on, we assume the force has been appropriately regularized.

Fixed Point Characterization

To prove the convergence of the point particle dynamics to the Vlasov-Poisson System, we have to use a more sophisticated characterization of the dynamics. Let μ_t be a time dependent probability measure so that

$$\int \varphi d\mu_t$$

is continuous in t for every bounded, continuous φ. With the modified force $G(x,y)$, we can replace our system of PDEs by one equation:

$$\partial_t \mu_t + v \cdot \nabla_q \mu_t + \left(\iint G(q,q') d\mu_t(v,q') \right) \cdot \nabla_v \mu_t = 0.$$

Associated to this PDE is a vector field on $\mathbb{R}^6$:

$$V[\mu_\cdot](t,v,q) = \begin{bmatrix} \iint G(q,q') d\mu_t(v,q') \\ v \end{bmatrix}.$$

Associated to the vector field $V[\mu_\cdot]$ is a flow on $\mathbb{R}^6$, $T_{t,0}[\mu_\cdot](v, q)$. This flow is characterized by

$$T_{t,0}[\mu_\cdot](\mathring{v}, \mathring{q}) = (v(t), q(t))$$

where $(v(t), q(t))$ satisfies the first-order system of ODEs

$$\begin{aligned} \dot{v}(t) &= \iint G(q(t), q')d\mu_t(v, q') \\ \dot{q}(t) &= v(t) \end{aligned}$$

and $(v(0), q(0)) = (\mathring{v}, \mathring{q})$. We can now write our PDE very succinctly by

$$\mu_t(v, q) = \mu_0 \circ T_{0,t}[\mu_\cdot](v, q).$$

We have finally laid all the necessary groundwork to prove our main theorem. Our proof rests on proving an inequality of the type

$$d_{bL^*}(\mu_t, \nu_t) \leq C(t) d_{bL^*}(\mu_0, \nu_0)$$

where $C(t) > 0$ and both μ_t and ν_t satisfy the fixed point equation we developed above. To begin, we list two vital properties satisfied by our vector field $V[\mu_\cdot]$:

We have finally laid all the necessary groundwork to prove our main theorem. Our proof rests on proving an inequality of the type

$$d_{bL^*}(\mu_t, \nu_t) \leq C(t) d_{bL^*}(\mu_0, \nu_0)$$

where $C(t) > 0$ and both μ_t and ν_t satisfy the fixed point equation we developed above. To begin, we list two vital properties satisfied by our vector field $V[\mu_\cdot]$:

- $V[\mu_\cdot](t, v, q)$ is continuous in t and Lipschitz continuous in (v, q) with Lipschitz constant L (independent of t).

We have finally laid all the necessary groundwork to prove our main theorem. Our proof rests on proving an inequality of the type

$$d_{bL^*}(\mu_t, \nu_t) \leq C(t) d_{bL^*}(\mu_0, \nu_0)$$

where $C(t) > 0$ and both μ_t and ν_t satisfy the fixed point equation we developed above. To begin, we list two vital properties satisfied by our vector field $V[\mu.]$:

- $V[\mu.](t, v, q)$ is continuous in t and Lipschitz continuous in (v, q) with Lipschitz constant L (independent of t).
- The mapping $\mu. \to V[\mu.]$ satisfies

$$\iint |V[\mu.](t, v, q) - V[\nu.](t, v, q)| \, d\mu_t(v, q) \leq K d_{bL^*}(\mu_t, \nu_t)$$

for some constant $K > 0$.

We have finally laid all the necessary groundwork to prove our main theorem. Our proof rests on proving an inequality of the type

$$d_{bL^*}(\mu_t, \nu_t) \leq C(t) d_{bL^*}(\mu_0, \nu_0)$$

where $C(t) > 0$ and both μ_t and ν_t satisfy the fixed point equation we developed above. To begin, we list two vital properties satisfied by our vector field $V[\mu_\cdot]$:

- $V[\mu_\cdot](t, v, q)$ is continuous in t and Lipschitz continuous in (v, q) with Lipschitz constant L (independent of t).
- The mapping $\mu_\cdot \to V[\mu_\cdot]$ satisfies

  $$\iint |V[\mu_\cdot](t, v, q) - V[\nu_\cdot](t, v, q)| \, d\mu_t(v, q) \leq K d_{bL^*}(\mu_t, \nu_t)$$

  for some constant $K > 0$.
- NOTE that both K and L depend on our regularization!

$$\begin{aligned} d_{bL^*}(\mu_t,\nu_t) &= d_{bL^*}(\mu_0\circ T_{0,t}[\mu.],\nu_0\circ T_{0,t}[\nu.]) \\ &\le d_{bL^*}(\mu_0\circ T_{0,t}[\mu.],\nu_0\circ T_{0,t}[\mu.]) \\ &\quad + d_{bL^*}(\nu_0\circ T_{0,t}[\mu.],\nu_0\circ T_{0,t}[\nu.]) \end{aligned}$$

The first term is fairly easy to estimate:

$$\begin{aligned} d_{bL^*}(\mu_0\circ T_{0,t}[\mu.],&\nu_0\circ T_{0,t}[\mu.]) \\ &= \sup_{\varphi\in D}\left|\iint \varphi d(\mu_0\circ T_{0,t}[\mu.]-\nu_0\circ T_{0,t}[\mu.])\right| \\ &= \sup_{\varphi\in D}\left|\iint \varphi\circ T_{t,0}[\mu.]d(\mu_0-\nu_0)\right| \\ &= e^{Lt}\sup_{\varphi\in D}\left|\iint e^{-Lt}\varphi\circ T_{t,0}[\mu.]d(\mu_0-\nu_0)\right| \\ &\le e^{Lt}d_{bL^*}(\mu_0,\nu_0) \end{aligned}$$

The second term requires more work.

$$\begin{aligned} d_{bL^*}(\nu_0 \circ T_{0,t}[\mu_\cdot], \nu_0 \circ T_{0,t}[\nu_\cdot]) \\ &= \sup_{\varphi \in D} \left| \iint \varphi d(\nu_0 \circ T_{0,t}[\mu_\cdot] - \nu_0 \circ T_{0,t}[\nu_\cdot]) \right| \\ &= \sup_{\varphi \in D} \left| \iint (\varphi \circ T_{t,0}[\mu_\cdot] - \varphi \circ T_{t,0}[\nu_\cdot]) d\nu_0 \right| \\ &\leq \iint \left| T_{t,0}[\mu_\cdot] - T_{t,0}[\nu_\cdot] \right| d\nu_0 \end{aligned}$$

Let $\lambda(t) = \iint \left| T_{t,0}[\mu_\cdot] - T_{t,0}[\nu_\cdot] \right| d\nu_0$.

Estimating $\lambda(t)$ requires using the vector field V associated to the flow T. With a little effort, we find that

$$\begin{aligned}
\lambda(t) &\leq \int_0^t \left(\iint \left| V[\mu.](\tau, v, q) - V[\nu.](\tau, v, q) \right| d\nu_\tau(v, q) \right) d\tau \\
&\quad + L \int_0^t \iint \left| T_{\tau,0}[\mu.](v, q) - T_{\tau,0}[\nu.](v, q) \right| d\nu_0(v, q) d\tau \\
&\leq K \int_0^t d_{bL^*}(\mu_\tau, \nu_\tau) d\tau + L \int_0^t \lambda(\tau) d\tau
\end{aligned}$$

The corollary to Gronwall's Inequality gives us that

$$\lambda(t) \leq Ke^{Lt} \int_0^t d_{bL^*}(\mu_\tau, \nu_\tau) e^{-L\tau} d\tau.$$

Putting this together with our first estimate

$$d_{bL^*}(\mu_t, \nu_t) \leq e^{Lt} d_{bL^*}(\mu_0, \nu_0) + Ke^{Lt} \int_0^t d_{bL^*}(\mu_\tau, \nu_\tau) e^{-L\tau} d\tau.$$

A final application of Gronwall's Inequality gives us the desired inequality

$$d_{bL^*}(\mu_t, \nu_t) \leq e^{(K+L)t} d_{bL^*}(\mu_0, \nu_0).$$

Conclusions

Applying these results to our specific case gives

$$d_{bL^*}(\Delta_t^N, f_t) \quad \leq \quad e^{(K+L)t} d_{bL^*}(\Delta_0^N, f_0).$$

Conclusions

Applying these results to our specific case gives

$$d_{bL^*}(\Delta_t^N, f_t) \leq e^{(K+L)t} d_{bL^*}(\Delta_0^N, f_0).$$

- With this inequality, convergence of the initial data ensures convergence at all later times.

Conclusions

Applying these results to our specific case gives

$$d_{bL^*}(\Delta_t^N, f_t) \quad \leq \quad e^{(K+L)t} d_{bL^*}(\Delta_0^N, f_0).$$

- With this inequality, convergence of the initial data ensures convergence at all later times.
- Moreover, for any physical system (with a finite number of particles), this inequality gives an idea of how long we can expect the fluid model to well approximate the particle dynamics.

- Neunzert, H.,
  An Introduction to the Nonlinear Boltzmann-Vlasov Equation,
  Lecture Notes in Mathematics, Springer-Verlag, Berlin (1984).
- Rudin, W.,
  Functional Analysis: Second Edition,
  McGraw-Hill, Boston (1991).
- Strichartz, R.,
  A Guide to Distribution Theory and Fourier Transforms,
  World Scientific Publishing Co., Danvers, MA (2003).