Stochastic Control Theory & Automated Market Making

Nov 19, 2010

Doug Borden
Electronic Trading Group
Knight Equity Markets
dborden@knight.com

Knight

Contents

Introduction

Who am I?

I work in the Electronic Trading Group (ETG) at Knight Capital. In my career I previously worked at Credit Suisse and at Goldman Sachs. I started my career in derivatives, and later switched to quantitative cash trading. At Knight I work to ensure optimal execution across our electronic strategies.

Who is Knight?

- Knight is a Broker-Dealer originally established to cater to the retail community
- Knight is the leading source of off-exchange liquidity in U.S. equities across all market segments
- Knight provides market making and agency-based trading in U.S., European and Asian equities, ADRs, ETFs, futures and options
- Knight is the largest U.S. market-maker, trading in more than 19,000 U.S. Equities
- Knight provides connectivity to more than 100 external market centers worldwide, including exchanges, ECNs, ATSs, dark pools, ATFs, MTFs and broker-dealers
- In 2009, Knight traded approximately 2.5 trillion shares and executed more than 980 million trades, an average of 600,000 trades per hour
- Knight is #1 in shares traded of Listed securities with 19.0% market share*
- Knight is #1 in shares traded on NASDAQ Capital Market, Global Market, and Global Select Market securities with a combined market share of 21.9%*
- Knight is #1 in shares traded of Bulletin Board and Pink Sheet securities with 85.7% market share*
- Knight is #1 in shares traded of S&P 500 securities with a 15.3% market share*
- Knight also trades in Fixed Income, Currencies & Commodities, and offers Capital Markets and Investment Banking services as well.

*Autex, 2010 YTD

High Frequency Trading Decisions

How do High Frequency Trading decisions differ from trading decisions associated with "Investing"?

	Long-dated Trading	**High Frequency Trading**
Holding Period	Weeks to Years	Seconds to Minutes
Expected Returns	1% -- 10% per position >> typical bid-offer spread	.01% -- .10% per position ~ typical bid-offer spread
Trading Style	Liquidity Taking	Liquidity Providing

I'm going to be discussing high frequency trading in the context of providing liquidity to the markets – i.e. voluntarily placing bids and offers on a large number of stocks into public and semi-public trading venues (exchanges, ECN's, dark pools, etc.).

When engaging in such activity, the high frequency trader hopes to get paid the bid-ask spread in return for suffering negative selection (i.e. he or she is more likely to be a buyer in a falling market and a seller in a rising market).

Generally the difference between profit and loss is very thin, and making real-time decisions on size and price is a difficult and delicate affair.

Parametric Approach

Imagine that for every stock in my universe, at all times, I have a short-dated forecast for the order-book. How do I use this information to construct a profitable trading strategy?

Toy Model

Here is a simple rule for a market-making trading strategy:

1. **Size: 100 shares to buy; 100 shares to sell**
2. **Price: BuyPrice = Mid-Market + A x Forecast – B x MCR – Average half-spread**

   SellPrice = Mid-Market + A x Forecast – B x MCR + Average half-spread

where MCR is the stock's marginal contribution to my portfolio risk.

I then have two parameters, A and B, which I determine by backtesting my strategy.

Generally this is a sound approach, albeit in practice the models will have many, *many* more parameters than the two I've specified here.

Parametric Approach (cont.)

What are the strengths and weaknesses of the parametric approach?

	Strength	Weakness
Rules are intuitive	✓	
Optimization tends to be stable	✓	
Edge cases handled poorly		✓
Rule-sets can grow to be very complex		✓
Not adaptive to changes in market structure		✓

Could Stochastic Control Theory offer a better approach?

Not a new idea:

“Applications of Mathematical Control Theory to Finance: Modeling Simple Dynamic Cash Balance Problems”, S. Sethi and G. L. Thompson (1970), *Journal of Financial and Quantitative Analysis*, **5**, pp. 381-394.

“Optimal Investment Models and Risk-Sensitive Stochastic Control”, W. H. Fleming (1995), *IMA Volume of Mathematical Finance*, **65**, pp. 35-45.

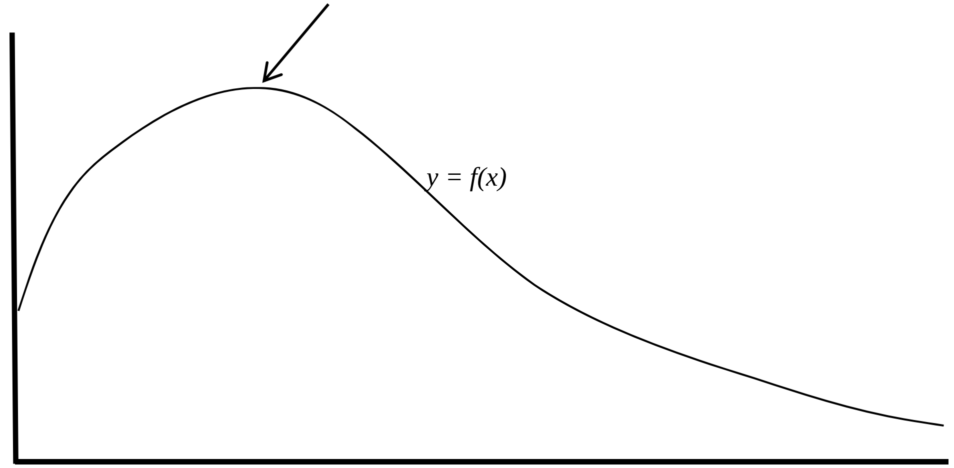
“An Application of Stochastic Control Theory to Financial Economics”, W. H. Fleming and T. Pang (2003), *SIAM Journal on Control and Optimization*, **43**, pp. 502-531.

… and many more.

Freshman Calculus

Given a function $y = f(\mathbf{x})$, we can find the values of $\mathbf{x}$ where y assumes its maximum or minimum by finding solutions to the algebraic equation:

$$\nabla f\Big|_{\mathbf{x}=\mathbf{x}_o} = 0$$

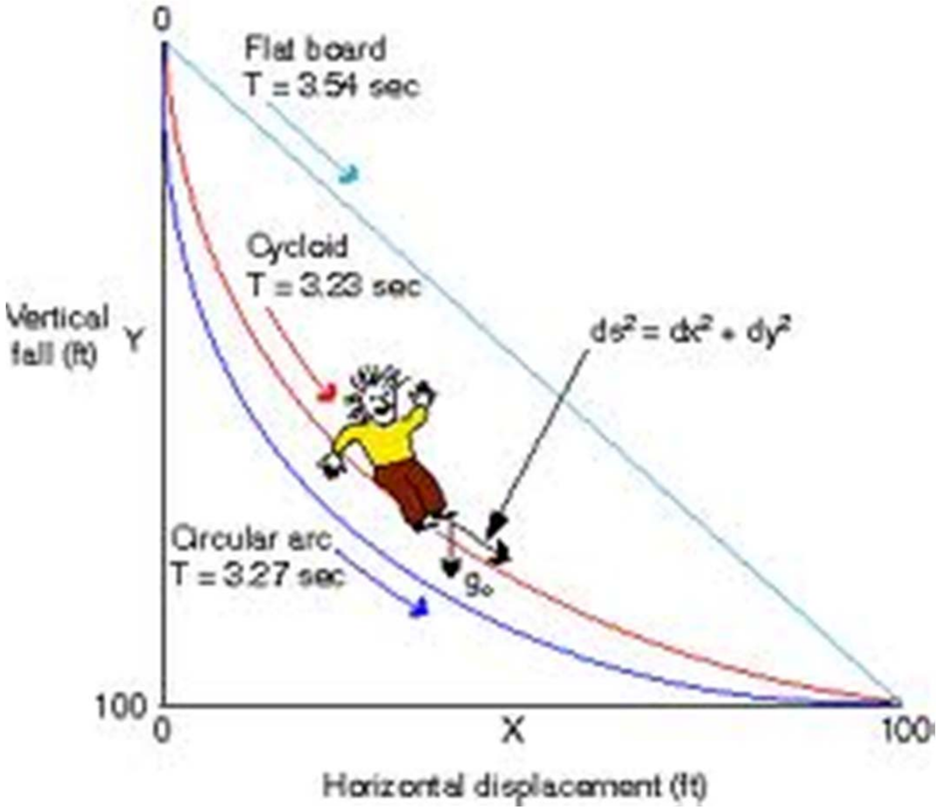


Sophomore Calculus

Given a functional $F[\mathbf{y}] = \int L(\mathbf{y}(x), \mathbf{y}'(x))dx$, we can find the function $\mathbf{y}^*(x)$ where $F[\mathbf{y}]$ assumes its maximum or minimum by finding solutions to the differential equation:

$$\nabla_y L - \frac{d}{dx}\nabla_{y'} L\Big|_{y(x)=y^*(x)} = 0$$

(The Euler-Lagrange Eqns)



Example The brachistochrone problem: Given a starting point *"a"* and an ending point *"b"* (with *a* at a higher elevation than *b*), what shape shall we give a slide such that a frictionless child would slide from *a* to *b* in the least amount of time.

Senior Calculus

Given $\boldsymbol{x}$ a vector of state variables, $\boldsymbol{u}$ a vector of controls, known functions $\boldsymbol{F}$, $\boldsymbol{G}$, L and B, and R a functional:

$$d\boldsymbol{x} = \boldsymbol{F}[\boldsymbol{x}(t), \boldsymbol{u}(\boldsymbol{x}, t), t]dt + \boldsymbol{G}[\boldsymbol{x}(t), \boldsymbol{u}(\boldsymbol{x}, t), t]dz$$

$$R[\boldsymbol{u}] = \int_0^T L[\boldsymbol{x}(t), \boldsymbol{u}(\boldsymbol{x}, t), t]dt + B[\boldsymbol{x}(T), T]$$

The optimal controls $\boldsymbol{u}^*$, at which R attains its extremum, are given by:

$$\nabla_{\boldsymbol{u}}\left[\nabla_{\boldsymbol{x}} H \cdot \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{u}, t) + \tfrac{1}{2}\boldsymbol{G}^T(\boldsymbol{x}, \boldsymbol{u}, t) \cdot \nabla_{\boldsymbol{x}}^2 H \cdot \boldsymbol{G}(\boldsymbol{x}, \boldsymbol{u}, t) + L(\boldsymbol{x}, \boldsymbol{u}, t)\right]_{\boldsymbol{u}=\boldsymbol{u}^*} = 0$$

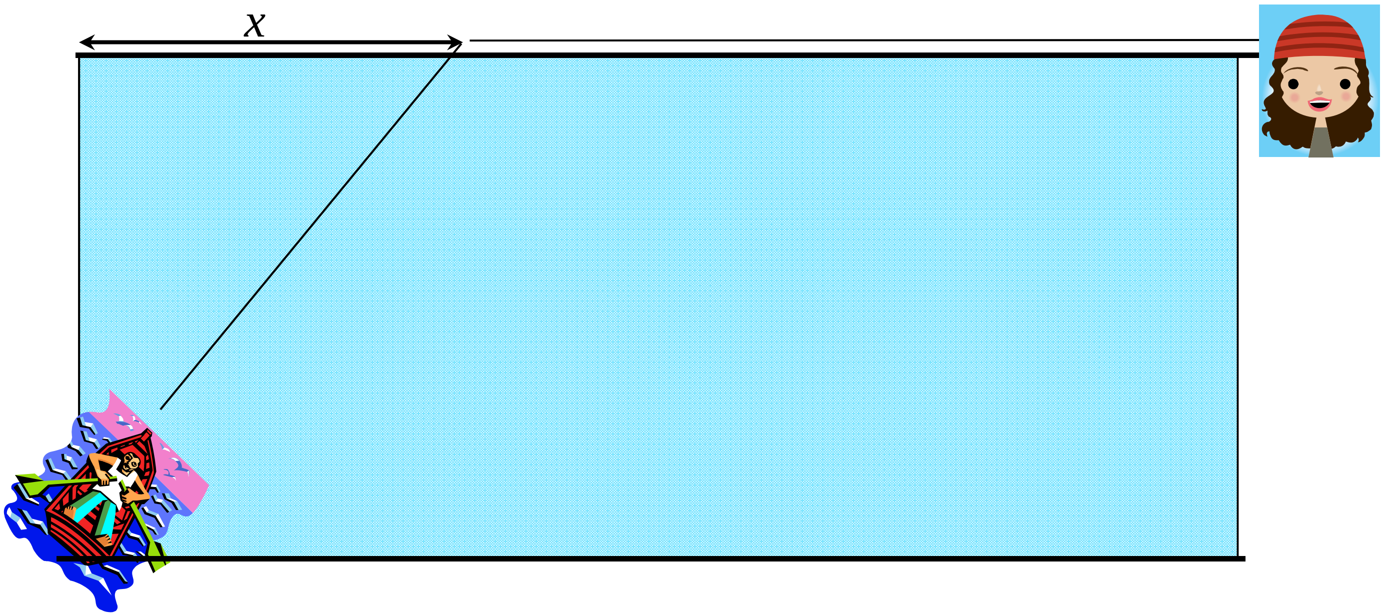
Where *"H"* is called the Value Function and satisfies the differential equation:

$$\frac{\partial H}{\partial t} + \nabla_{\boldsymbol{x}} H \cdot \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{u}^*, t) + \tfrac{1}{2}\boldsymbol{G}^T(\boldsymbol{x}, \boldsymbol{u}^*, t) \cdot \nabla_{\boldsymbol{x}}^2 H \cdot \boldsymbol{G}(\boldsymbol{x}, \boldsymbol{u}^*, t) + L(\boldsymbol{x}, \boldsymbol{u}^*, t) = 0$$

with boundary condition: $H(\boldsymbol{x}, T) = B(\boldsymbol{x}, T)$

The above equations are known as the **Hamilton-Jacobi-Bellman Equations**.

The problem of the guy and the rowboat:

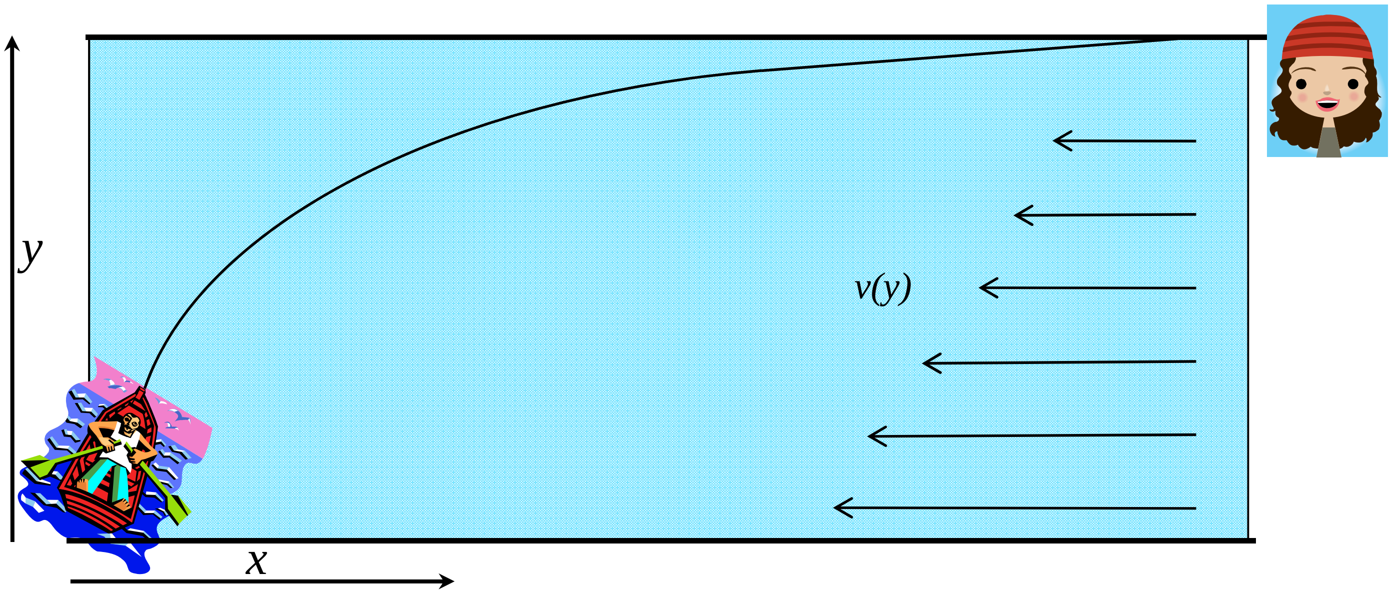


Jack wants to get to Jill as fast as he can, who is 2 miles down-river from him, but he needs to cross the ½-mile wide river to get to her. He can row at 3mph, but once across he can run at 9mph along the bank. Toward which point along the bank should Jack row so as to reach Jill in the smallest amount of time *"T"*?

$$T = \sqrt{1/4 + x^2} / 3 + (2 - x) / 9 \qquad dT/dx|_{x=x_o} = 0 \Rightarrow x_o = 0.1767 \text{ miles}$$

We say that the optimal *point "x_o"* is a solution to the algebraic equation *dTime/dx = 0*.

What if there is a variable current with a velocity *v(y)* [and the shore is too rocky to run along]?



Now we need to find the optimal ***path*** *$y_o(x)$* for Jack to row. For any path *y(x)*, the time it takes Jack to reach Jill is given by:

$$T = \int_0^2 \sqrt{\frac{(1+y'^2)^{3/2}}{\left(9+v(y)^2\right)\sqrt{1+y'^2}-6v(y)}}\,dx \equiv \int_0^2 L(y,y')\,dx$$

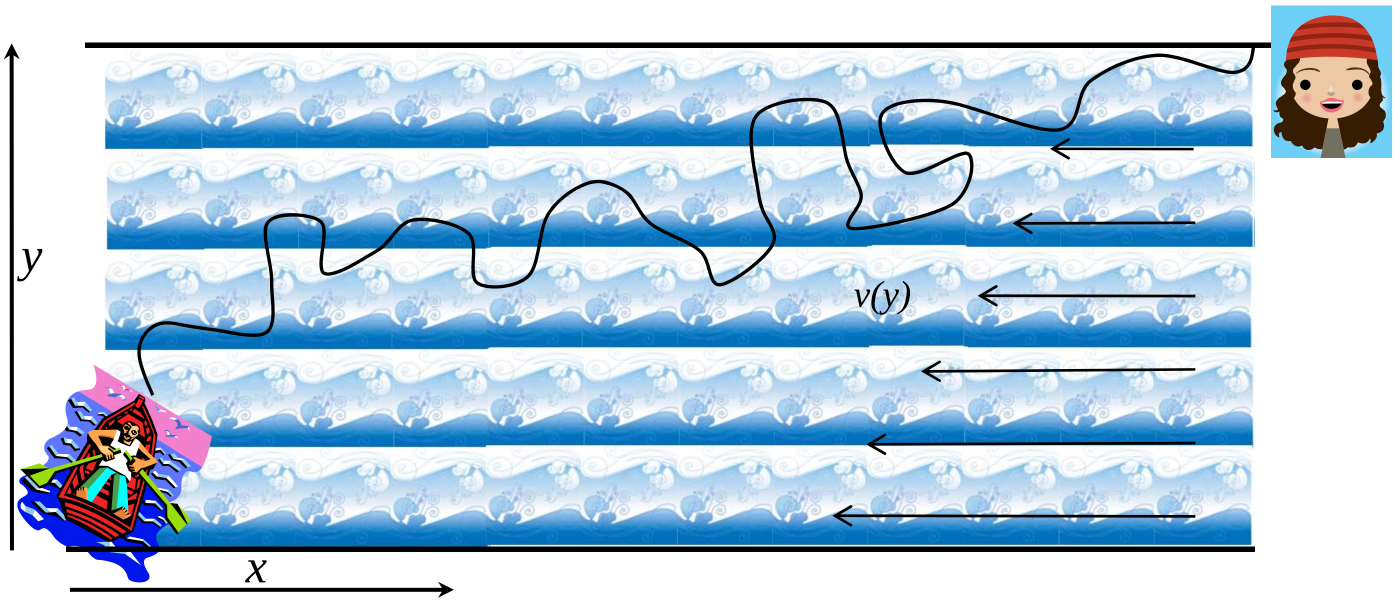
and the *optimal path $y_o(x)$* must satisfy the Euler-Lagrange Equation:

$$\left.\frac{d}{dx}\frac{\partial L}{\partial y'} - \frac{\partial L}{\partial y}\right|_{y(x)=y_o(x)} = 0$$

In this case we say the optimal path *$y_o(x)$* is a solution to the differential equation $\frac{d}{dx}\frac{\partial L}{\partial y'} - \frac{\partial L}{\partial y} = 0$

Review of Variational Calculus (cont.)

Now let's add turbulence to the river, so that as he rows across, Jack's rowboat is tossed about by the rapids.



How do we help Jack get to Jill in this situation? Jack's path to Jill is no longer deterministic – though his boat's motion is partially determined by the direction in which Jack rows, there is also a random component to the motion due to the turbulence. In a case such as this we need to provide Jack with an ***optimal policy*** which he can follow – instructions for Jack which will optimize his path to Jill no matter where in the river he finds himself. This policy will take as inputs Jack's current position and velocity, and will output the optimal direction that Jack should row. Finding optimal *policies* is the job of Stochastic Control Theory and the HJB equations.

Knight

Toy Model

Let's consider the following simplified model for Stock ABC and the market in which it trades:

- ABC trades continuously (no quantized fills).
- The mid-market price of ABC follows algebraic Brownian motion (*μ(t) known*).

$$dS = \mu(t)dt + \sigma dz$$

- The spread of ABC is constant with a width 2δ (like, say, Citigroup).
- Market participants can place passive **limit** orders to buy at the bid or to sell at the ask, or they may place active **market** orders to buy at the ask or sell on the bid. Note that limit orders "earn" spread and market orders "pay" spread.
- Market orders (both buys and sells) are filled at a constant rate $f_o > v_o$.
- Limit orders to buy and sell are filled at rates given by:

$$dv_s = \beta(v_o - v_s)dt + \varsigma dz$$

$$dv_b = \beta(v_o - v_b)dt - \varsigma dz$$

- There is no market impact.

As market makers, what do we get to control? We can control four variables (at each time t):

1. Whether we place a limit order to buy → Let's define this as $\lambda_b(t)$ which takes values of either 0 or 1
2. Whether we place a limit order to sell → Let's define this as $\lambda_s(t)$ which takes values of either 0 or 1
3. Whether we place a market order to buy → Let's define this as $m_b(t)$ which takes values of either 0 or 1
4. Whether we place a market order to sell → Let's define this as $m_s(t)$ which takes values of either 0 or 1

Let's denote $N(t)$ the number of shares we hold in inventory at time t. Then $N(t)$ is given by:

$$dN = \left[-\lambda_s(t)v_s(t) + \lambda_b(t)v_b(t) - m_s(t)\,f_o + m_b(t)\,f_o\right]dt$$

What quantity do we want to extremize? We want to maximize our risk-adjusted PnL:

$$U = \int_0^T \left[\left(S(t)+\delta\right)\lambda_s(t)v_s(t) - \left(S(t)-\delta\right)\lambda_b(t)v_b(t) + \left(S(t)-\delta\right)m_s(t)\,f_o - \left(S(t)+\delta\right)m_b(t)\,f_o - \eta\sigma^2 N^2(t)\right]dt$$

with η our risk aversion coefficient.

So we have 4 state variables (S, v_s, v_b, N), each with an evolution equation, and 4 controls $(\lambda_b, \lambda_s, m_b, m_s)$. Plugging into the HJB equations …

… gives a mildly intimidating set of 2nd order partial differential equation for $H = H(S,v_s,v_b,N,t)$:

$$\frac{\partial H}{\partial t}+\mu(t)\frac{\partial H}{\partial S}+\beta(v_o-v_s)\frac{\partial H}{\partial v_s}+\beta(v_o-v_b)\frac{\partial H}{\partial v_b}-\left(\lambda_s v_s-\lambda_b v_b+m_s f_o-m_b f_o\right)\frac{\partial H}{\partial N}$$

$$+\tfrac{1}{2}\sigma^2\frac{\partial^2 H}{\partial S^2}+\tfrac{1}{2}\zeta^2\frac{\partial^2 H}{\partial v_s^2}+\tfrac{1}{2}\zeta^2\frac{\partial^2 H}{\partial v_b^2}+\sigma\zeta\frac{\partial^2 H}{\partial S\partial v_s}+\sigma\zeta\frac{\partial^2 H}{\partial S\partial v_b}+\zeta^2\frac{\partial^2 H}{\partial v_b\partial v_s}$$

$$+\left(S+\delta\right)\lambda_s v_s-\left(S-\delta\right)\lambda_b v_b+\left(S-\delta\right)m_s f_o-\left(S+\delta\right)m_b f_o-\eta\sigma^2 N^2=0$$

$$\nabla_{\lambda,m}\Bigg[\ \mu(t)\frac{\partial H}{\partial S}+\beta(v_o-v_s)\frac{\partial H}{\partial v_s}+\beta(v_o-v_b)\frac{\partial H}{\partial v_b}-\left(\lambda_s v_s-\lambda_b v_b+m_s f_o-m_b f_o\right)\frac{\partial H}{\partial N}$$

$$+\tfrac{1}{2}\sigma^2\frac{\partial^2 H}{\partial S^2}+\tfrac{1}{2}\zeta^2\frac{\partial^2 H}{\partial v_s^2}+\tfrac{1}{2}\zeta^2\frac{\partial^2 H}{\partial v_b^2}+\sigma\zeta\frac{\partial^2 H}{\partial S\partial v_s}+\sigma\zeta\frac{\partial^2 H}{\partial S\partial v_b}+\zeta^2\frac{\partial^2 H}{\partial v_b\partial v_s}$$

$$+\left(S+\delta\right)\lambda_s v_s-\left(S-\delta\right)\lambda_b v_b+\left(S-\delta\right)m_s f_o-\left(S+\delta\right)m_b f_o-\eta\sigma^2 N^2\ \Bigg]^*=0$$

***This notation should be taken to mean the extremum over ths λ's and m's, which in this case means a generalized Heavyside step function in the four state variables, H and ∇H**

Is this solvable?	Yes – over the past couple of decades, the fields of physics, engineering, communications, operations research, finance, etc. have provided the scientific community with a wealth of tools for solving differential equations like these.
What about speed?	Very difficult – most current techniques require ~ seconds to converge, which is 3 orders of magnitude too slow for high frequency trading decisions. But specialized hardware, pre-computing, and using recently computed solutions plus variations can speed things up considerably.
How well does it work?	Too soon to tell ….

- Stochastic Control Theory provides a rigorous framework for making decisions under conditions of uncertainty.

- High Frequency Trading decisions lend themselves to be cast in such a framework.

- Even the simplest of market models leads to very complicated differential equations.

- BUT: Physics, Engineering, Operations Research, and two decades of derivatives pricing all provide a wealth of tools for solving the resulting HJB equations.