POMS PRODUCTION AND OPERATIONS MANAGEMENT

PRODUCTION AND OPERATIONS MANAGEMENT POMS
Vol. 23, No. 11, November 2014, pp. 1875–1885
ISSN 1059-1478 | EISSN 1937-5956 | 14 | 2311 | 1875
DOI 10.1111/poms.12188

Of Physics and Factory Physics

Mark L. Spearman
Factory Physics, Inc, 3600 E. 29th Street, Bryan, Texas 77802, USA, spearman@factorphysics.com

The "Principle of Least Action" is the foundational principle of fundamental physics. Application of this principle to the supply chain naturally results in an "uncertainty principle" linking variation in production to variation in net-inventory. The model also provides an easy means of determining the stationary distribution of net-inventory for a variety of control strategies. The formalism results in a control strategy that outperforms commonly used control methods.

Key words: production-inventory models; Brownian motion; uncertainty
History: Received: May 2013; Accepted: September 2013 by Kalyan Singhal, after 1 revision.

1. Introduction

Factory Physics is a text (Hopp and Spearman 2011) that first appeared in 1996 and sought to provide a systematic framework to describe the behavior of manufacturing systems. Like "real" physics, factory physics is stated as a set of "laws" about the way production-inventory systems behave. Unlike "real" physics, the laws of factory physics do not follow from a single unifying framework.

Lord Rutherford, discoverer of the nucleus, once stated, "All science is either physics or stamp collecting," implying that the only true science was physics and all the others were simply cataloging various observations.[1] Alas, it appears that factory physics has progressed little beyond the stamp collecting stage.

The purpose of this study is to move toward a more general framework that will describe the behavior of production-inventory systems. By doing this, we may obtain further insights into how manufacturing systems behave and perhaps even suggest better ways to design and control them.

An important factory physics "law" deals with variability—the greater the variability the poorer the performance. The qualitative model presented in the text postulates that the only way that production and demand can be synchronized in the presence of variability is to employ some sort of "clutch" or "buffer" between the two and that there are only three possibilities: inventory, time, and/or capacity. The inventory buffer represents any inventory between production and demand. Likewise, the time buffer represents any time between the point when the demand first occurred and when it is finally satisfied. The capacity buffer is defined to be the difference between available capacity and the average demand.

A vertically integrated "configure-to-order" (CTO) operation provides a good example of how these three buffers interact. Such systems usually wait until a customer gives a, possibly, unique specification and then configures (or assembles) the final product to that specification. Consequently, the system employs both a time buffer between the CTO operation and the customer and an inventory buffer between the final assembly process and fabrication (or a supplier). In the later case, the assembly process represents "demand" and the fabrication (supplier) provides the "production."

However, if such a system were to use only an inventory buffer, it would require stock for each and every possible configuration of which there can be an enormous number. If it were to employ only a time buffer, the customer would have to wait not only the time for the desired item to be configured but also the time it would take to acquire the components. By using a combination, inventory of components, and time to configure, the system has neither excessive inventory nor lengthly wait times.

The capacity buffer affects both the inventory and time buffers. With a significant amount of capacity above the average demand, the fabrication time can be kept low. However, if there is not much extra capacity, this queue can become very long requiring a longer wait and typically more inventory.

So it is clear that these three buffers are interrelated. The question is how are they related and how can that relation be generalized.

1.1. This Study

This study is addressed to the operations researcher interested in supply chain applications. Because it will consider some fundamental models from physics, this study will necessarily involve some tutorial. This will assume that the reader is mathematically sophisticated, but has little or no knowledge of fundamental physics. We begin with a "basic" model of a production-inventory system that models a base stock system in the presence of demand variation. Using Itô calculus we derive a relation between the variability in the net-inventory and production. The form of this result is similar to a fundamental relation found in quantum physics which motivates the remainder of the article.

We then consider the fundamental model that underlies almost all branches of physics, from classical mechanics to quantum field theory—The Principle of Least Action, also known as Hamilton's Principle (see, e.g., Stevens 1995). Adding randomness and then applying a "potential" to control the "motion" of an inventory "particle," we arrive at the following conclusions:

1. The "buffers" mentioned above are linked in a "canonical" way using the Least Action formalism.
2. There is an "uncertainty principle" between production and net-inventory (i.e., as the variance of one decreases, the other must increase).
3. We can use a "potential" to set the control of a production-inventory system.
4. We can easily write down the final (steady-state) joint distribution of net-flow and net-inventory in terms of the system "Hamiltonian."
5. The model used in traditional MRP is a *special case* of the Least Action model. Consequently, this model cannot perform worse than the traditional model and will, in many cases, perform better.

2. A Basic Production-Inventory Model

We start with a simple general production-inventory model that can be written as

$$Q(t+\Delta t) = Q(t) + X(t) - D(t), \tag{1}$$

where Q represents net-inventory, X represents production during period t, a decision variable, and D represents demand, usually random, with a mean of $\lambda\Delta t$ during the same period. In general, we can define

$$D(t) = \lambda\Delta t + W(\Delta t),$$

where λ is the mean of the demand and W is a random variable with zero mean and a variance that depends on Δt. We also define,

$$X(t) = (\lambda + g(Q(t))\Delta t \tag{2}$$

$$= R(t)\Delta t, \tag{3}$$

where g is some real function of the net-inventory. This basic model exhibits the behavior of many production systems including a base stock system, a reorder point/reorder quantity system, and even a time-phased reorder point system that seeks to keep inventory at a given level (safety stock).

A common version of this model is $g(Q) = -g(Q(t) - Q_0)$, where Q_0 is a base stock level (here set equal to 0 without loss of generality) and g is a constant and is usually set to 1. This model can also represent a time-phased reorder point (i.e., material requirements planning) system with a lot-for-lot order-sizing rule. We can model the transient and steady-state behavior of this system using a stochastic differential equations (SDEs). See, e.g., Oksendal (1995) for a complete treatment.

2.1. Formulation Using Stochastic Differential Equations

In many real-world cases we observe what is known as "quadratic variation" that implies that the variance of demand will increase linearly in time,

$$\lim_{t\to\infty} \frac{\mathrm{var}(Q(t))}{t} = \psi^2\lambda$$

Where ψ^2 is the variance to mean ratio of the demand. Then the SDE of the base stock model (with a target of zero) will be,

$$\begin{aligned} dQ &= -gQdt + \psi\sqrt{\lambda}dW \\ Q(0) &= Q_0 \end{aligned} \tag{4}$$

and is an example of the Ornstein-Uhlenbeck process whose solution is,

$$Q(t) = Q_0 e^{-gt} + \psi\sqrt{\lambda}\int_0^t e^{-g(t-s)}dW \tag{5}$$

The steady-state mean and variance are

$$\begin{aligned} \langle Q \rangle &= 0 \\ \mathrm{var}(Q) &= \frac{\psi^2\lambda}{2g}. \end{aligned} \tag{6}$$

The variance of production required to keep the variance of Q at the level given above will be,

$$\begin{aligned} \mathrm{var}(R) &= g^2\mathrm{var}(Q) \\ &= g\frac{\psi^2\lambda}{2}. \end{aligned} \tag{7}$$

If there is inherent variation in production, then (7) will be a lower bound so that

$$\sigma_R\sigma_Q \geq \frac{\psi^2\lambda}{2}. \tag{8}$$

The implication is that if we want tighter control on inventory, we must allow production to vary more. This accommodation requires extra capacity above and beyond the average demand to be available at all times so it will be available when it is needed to bring up the inventory to the desired target. This translates to a larger capacity buffer. On the other hand, if we allow more variation in net-inventory, we can live with a smaller capacity buffer. Of course, more variation in inventory implies more inventory and/or backorders, on average. These interactions are examples of the basic trade-off between the buffers as described in the introduction.

3. Relationship to Physics

The inequality (8) is reminiscent of a fundamental relation in quantum mechanics known as the Heisenberg Uncertainty Principle,

$$\sigma_p \sigma_q \geq \hbar/2,$$

where p and q are the momentum and position, respectively and $\hbar$ is the reduced Planck constant. Such relations occur in physics only between special pairs of variables known as *canonical conjugates*. Such variables are related via the *Lagrangian* of a dynamical system. Since supply chains are dynamic in nature and given the relation in (8), it seems reasonable to write the Lagrangian of a supply chain and see if we can model a real supply chain with it.

3.1. A Supply Chain Lagrangian

We will use the same notation as before, but with lower case variables.[2] We then define q to be a generalized coordinate denoting the net-inventory and its time derivative will be the net-flow, viz.

$$\dot{q} = x - d.$$

In mechanics, the motion of the particle determines its kinetic energy, whereas the potential energy determines how the particle will move. An example is the motion of the Earth within the gravitational potential of the Sun. If there were no potential, the Earth would move at a constant velocity in a straight line. But in the presence of a gravitational potential, the Earth orbits the Sun in a motion that forms an ellipse with the Sun at one of the foci.

Thus, control of net-inventory "particle" will be accomplished by a "potential" that generates a "force" on the particle to constrain its location. In addition to forces that control the particle, the potential also contains elements that generate the force resulting in the demand of the system. Fortunately, the potential is a scaler so it is easy to add these components. Also, as in physics, it makes things more convenient if the potential is a function of only the coordinate and not time or the net-flow, $V(q)$.

The Lagrangian[3] is defined to be the difference between the kinetic energy and the potential energy. Typically, the kinetic energy involves a mass term that will not be included here. Mass in physics is related to inertia and to gravity. While inventory particles are not affected by gravity they will have inertia in many situations such as when production rates are not continuously varied but set periodically. As we will see, our formulation matches these situations exactly—without a mass term. Consequently, to avoid unnecessary notation we will set all the masses to 1 and drop them from the subsequent development. While massless formulations are not common in physics, they do occur (e.g., a Lagrangian in quantum field theory can be written such that there is no explicit mass term). However, this will mean that the formulation developed here will *not* be appropriate for situations in which the control response *can* be varied instantaneously and continuously.

Thus, the Lagrangian involves the kinetic energy, $\dot{q}/2$, and the potential energy, V(q), and is written as

$$L(q, \dot{q}, t) = \frac{1}{2}\dot{q}^2 - V(q).$$

Then, for *any* generalized coordinate, a corresponding "momentum" is defined as

$$p = \frac{\partial L(q, \dot{q})}{\partial \dot{q}}.$$

Defined in this manner, the momentum becomes the *canonical conjugate* of the generalized position, q, and in this case is equal to the net-flow,

$$p = \dot{q} = x - d. \tag{9}$$

3.2. The Hamiltonian and the Principle of Least Action

The overarching principle that governs the dynamics of the system is governed by the Principle of Least Action. Not only does this principle apply to mechanics, but it also applies to electricity and magnetism, optics, thermodynamics, and even quantum physics. In addition to physics, the Least Action principle has been applied to economics (Samuelson 1970), control theory (Sethi and Thompson 2000), medicine (Burns et al. 2012), biology (Kaila and Annila 2008, Samuelson 1974), and even robotics (De Sapio et al. 2007). The Least Action formalism is the underlying framework of the Standard Model of particle physics. As such, it has been spectacularly successful in not only explaining observed phenomena, but in also predicting new phenomena and new particles. In the summer of 2012, we saw the first observations of the Higgs boson (ATLAS 2012) that was predicted by the Standard Model almost 50 years earlier for which the Nobel Prize in physics was awarded in October of 2013.

Mathematically, Hamilton's Principle is written as,

$$\delta S = \delta \int_{t_i}^{t_f} L(q, \dot{q}, t) dt = 0,$$

where the δ indicates an infinitesimal variation in the path. This implies that the integral in any natural process will result in a *stationary* path, either a minimum or a maximum. Physical processes result in a minimum value.

Using the calculus of variations to find the Lagrangian function that minimizes the integral results in a necessary condition known as the *Euler–Lagrange equation* (see, e.g., Stevens 1995 for a derivation).

$$\frac{d}{dt}\left(\frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}}\right) - \frac{\partial L(q, \dot{q}, t)}{\partial q} = 0. \qquad (10)$$

In classical dynamics, force is defined as the time derivative of momentum, $\dot{p}$, we see from equation (10),

$$\dot{p} = \frac{\partial L(q, \dot{q})}{\partial q}.$$

Thus, $\dot{p}$ (or the "force") will be the change in the net-flow per unit time. If the mean demand remains constant, this change will occur in the production rate. Thus, the potential provides a means to describe the control used in the production-inventory system.

Upon reflection, the Principle of Least Action is astounding. If a ball is to roll down a hill, it will choose a path that minimizes the above integral. The problem here is not to simply find a point that minimizes a function, but to find a *function* that minimizes an integral. Given the difficulty of the former, how does a simple physical system accomplish the latter? Nonetheless, it does and the principle of least action leads directly to many laws of physics including Newton's laws of motion, Schrödinger's equation, relativistic electrodynamics, and quantum field theory.

A variant of the Least Action principle known as the "Maximum Principle" has been used in control theory for many years (see, e.g., Boltyanskii 1958). For an excellent and more recent exposition, see Chapter 2 in Sethi and Thompson (2000). Instead of being minimized, an integral similar to the action integral is maximized using dynamic programming methods. Maximization is appropriate as it represents the profit of the control policy and is in keeping with the action principle which seeks an *extremum* be it a minimum or a maximum. They then derive a generalized Hamilton–Jacobi equation and then use the construct to maximize the profit of a control trajectory. The "adjoint" variables used to denote marginal returns operate in the same way as canonical conjugates. However, the goal of this method is to determine optimal policies and not to model the particular dynamics of a given system.

Although the Lagrangian is written in terms of q and $\dot{q}$, we would like to be able to describe the system in terms of net-inventory, q, and net-flow, p. This can be done by applying a *Legendre transformation* to the Lagrangian resulting in a new function of q and p known as the *Hamiltonian*,

$$\mathcal{H}(q, p, t) = \frac{p^2(t)}{2} + V(q) \qquad (11)$$

4. Randomness and the Hamiltonian

While classical mechanics can conceivably model any system with an arbitrary number of particles, it can practically only be used to analyze relatively small systems. The study of thermodynamics allows us to study systems with a huge number of particles (on the order of 10^{23}). We will find a branch of thermodynamics called *statistical* physics most useful in that it provides us with a means to study random systems (although statistical physics is not really random at all).

We will follow the development of statistical physics and define a *phase space* of $2N$ dimensions (i.e., N ordered pairs, (q_i,p_i) representing the location and momentum of each particle i), where N is the number of particles. Each point in the phase space is known as a *microstate* with an associated probability and energy. Collections of microstates that have the same energy are called *macrostates* and we expect to find many microstates within a given macrostate. For example, the position and momentum of every molecule of the air in the room you are reading this article represents one microstate. The temperature and pressure of the room define a macrostate. Clearly, there are *many* microstates in this macrostate!

The central assumption of statistical physics is that all states of the system with the same energy have the same probability of occurring.

4.1. Liouville's Theorem

We define the density function of the phase space as $\rho(q,p;t)$ such that the expected value of a real function, $\langle g(t) \rangle$, is given by

$$\langle g(t) \rangle = \frac{\int_\Omega g(q, p) \rho(q, p; t) d\Omega}{\int_\Omega \rho(q, p; t) d\Omega},$$

where Ω represents the set of all states. One of the fundamental results of statistical physics is

$$\frac{d\rho}{dt} = 0,$$

which implies that a given volume in phase space may change shape during its evolution, but the volume will remain constant (see, e.g., Kardar 2007, for a proof). The result is

$$\frac{d\rho}{dt} = \frac{\partial\rho}{\partial t} + \frac{\partial\rho}{\partial q}\dot{q} - \frac{\partial\rho}{\partial p}\dot{p} = 0$$

or

$$\frac{\partial\rho}{\partial t} + \frac{\partial\rho}{\partial q}\frac{\partial\mathcal{H}}{\partial p} - \frac{\partial\rho}{\partial p}\frac{\partial\mathcal{H}}{\partial q} = 0, \tag{12}$$

which is known as "Liouville's Theorem" (see, e.g., Stevens 1995). If the system is in equilibrium, then $\frac{\partial\rho}{\partial t} = 0$ and so the second and third terms must also sum to zero. One way to achieve this is to assume that ρ is independent of q and p. A more general solution is

$$\rho(q,p) = C\exp\left(-\frac{\mathcal{H}(q,p)}{kT}\right), \tag{13}$$

which can easily be verified using (12) with $\frac{\partial\rho}{\partial t} = 0$. Note that kT is the energy of the system having "temperature" T so that k, Boltzmann's constant, provides the proportionality constant between energy and temperature.

Thus, the steady-state distribution of the net-inventory and the net flow in a production-inventory system described by Hamiltonian $\mathcal{H}$ can be written down by simply knowing the Hamiltonian and the final *temperature* of the system. So the question remains as to what is the "temperature" of a supply chain?

5. Modeling the Supply Chain

We now apply these concepts to a supply chain. We begin by considering models for demand. Then, we add potential terms to represent the control imposed.

We now postulate a Hamiltonian formulation for supply chain dynamics using the above developments.

5.1. A Hamiltonian Model for Demand

We first add a random force to represent the fluctuations in demand. These fluctuations are similar to what would be seen if a set of many small particles were continuously striking a larger particle in one dimension. However, in the presence of these many small particles there must also be some sort of *damping* process because, as Kubo (1966) explains, both the random fluctuations and the damping are caused by the same physical process.

Thus, in the presence of damping, we will require some sort of "force," representing the capacity less the demand of the system, pushing (or "pulling" if the enthusiasts insist) on the particle representing the net-inventory. If the capacity exceeds demand, inventory will continually increase and vice versa. This force can be represented by a linear potential, $V(q) = -F_0 q$, and the Hamiltonian will be

$$\mathcal{H} = \frac{p^2}{2} - F_0 q. \tag{14}$$

We will also need a damping term that results from a force that is proportional to the velocity of the particle acting in the opposite direction, $-\eta p$, where $1/\eta$ acts as a "time constant" indicating how quickly the system reaches the final value of net-flow.

To represent random fluctuations in demand, we add another force, $\beta F_r(t)$, where β is the "diffusion" coefficient and $F_r(t)$ is a stochastic process with mean zero and a variance that depends on t and is further specified below. Our equations of motion become,

$$\dot{p} = F_0 - \eta p + \beta F_r(t), \tag{15}$$

$$\dot{q} = p. \tag{16}$$

Despite these non-conservative forces, we can solve the resulting equations of motion. However, we would not expect for the energy of the system to remain constant, although, in steady state, it will achieve a fixed value.

5.2. Specifying the Stochastic Process

Interestingly, if we set $F_0 = 0$, equation (15) becomes *Langevin's equation* that was first proposed in 1908 to describe the dynamics of *physical* Brownian motion (Langevin 1908). Thus, it appears that Brownian motion arises naturally in the Hamiltonian formulation, although it is somewhat different from the Brownian motion processes found in the operations management literature.

For Brownian motion to be an appropriate model, we need only two conditions: that the path is continuous and that non-overlapping increments are probabilistically independent. The other conditions usually discussed (e.g., normality) can either be imposed or arise from these (see Harrison 1985 and the online Appendix of this study for a more complete discussion).

However, there are fundamental problems with this formulation. Because Brownian motion represents discrete jumps at a point in time, the model implies infinite velocities that are not possible.

Nonetheless, the model works very well and predicts results well within the experimental error.[4] The reason it works well is that whenever a "large" particle such as a pollen grain is suspended in a liquid, there are an incredible number of much smaller liquid molecules colliding with it each second. That the collisions are so numerous and are independent of each other results in normality via the Central Limit Theorem. For it to work well in a supply chain setting, we would need similar conditions—a large number of independent demands occurring during a typical time period. If the number of demands were less, we would need to increase the period size so that a significant number of demands (say, at least 10) occur in any period.

Equation (15) can be written as a SDE, viz.

$$dp = (F_0 - \eta p)dt + \beta dW, \tag{17}$$

$$p(0) = p_0,$$

where W represents a standard Brownian motion process (also known as a "Wiener" process) and β is the diffusion coefficient. Given the way W is defined, it has effective "dimensions" of $\sqrt{t}$ so that β has dimensions of $p/\sqrt{t}$. The solution to the above is obtained by the usual methods recognizing βdW as the differential of an "external" force.

$$p(t) = p_0 e^{-\eta t} + \frac{F_0}{\eta}\left(1 - e^{-\eta t}\right) + \beta \int_0^t e^{-\eta(t-s)} dW. \tag{18}$$

Again the last integral is a random variable whose expectation and variance may be found using the properties of Itô integrals. The expected value is,

$$\langle p(t)\rangle = \langle p_0\rangle e^{-\eta t} + \frac{F_0}{\eta}\left(1 - e^{-\eta t}\right), \tag{19}$$

so that the terminal velocity is F_0/η. This terminal velocity would correspond to the mean demand, λ, so that $\lambda = F_0/\eta$. The variance of p is found again using results from Itô calculus,

$$\mathrm{var}(p(t)) = \mathrm{var}(p_0)e^{-2\eta t} + \frac{\beta^2}{2\eta}\left(1 - e^{-2\eta t}\right), \tag{20}$$

so that the limiting variance is $\beta^2/(2\eta)$. If the potential does not depend on the momentum, then the limiting variance will alway converge to this value. In a controlled situation where the expected value of the momentum is zero, the limiting variance is also $\langle p^2\rangle$. Moreover, from the "Equipartition" theorem of statistical thermodynamics (see, e.g., Kardar 2007), we know

$$\langle p^2\rangle = kT = \frac{\beta^2}{2\eta}. \tag{21}$$

The position can be found by noting

$$q(t) = q_0 + \int_0^t p(s)ds,$$

and after some Itô calculus,

$$q(t) = q_0 + \frac{V_0}{\eta}(1 - e^{-\eta t}) + \frac{F_0}{\eta}\left(t - \frac{1}{\eta}(1 - e^{-\eta t})\right) + \frac{\beta}{\eta}\int_0^t (1 - e^{-\eta s})dW, \tag{22}$$

where q_0 and V_0 could be random variables and the last integral is again a random variable. The expected value is

$$\langle q\rangle = \langle q_0\rangle + \frac{\langle V_0\rangle}{\eta}(1 - e^{-\eta t}) + \frac{F_0}{\eta}\left(t - \frac{1}{\eta}(1 - e^{-\eta t})\right). \tag{23}$$

When q_0 and V_0 are constants, the variance of q is obtained by applying results from Itô calculus to (22).

$$\mathrm{var}(q(t)) = \frac{\beta^2}{\eta^2}\left(t - \frac{2}{\eta}(1 - e^{-\eta t}) + \frac{1}{2\eta}(1 - e^{-2\eta t})\right). \tag{24}$$

Note that as t goes to infinity, the variance becomes a linear function of time that can be related to the demand variance, viz.,

$$\lim_{t\to\infty} \frac{\mathrm{var}(q(t))}{t} = \frac{\beta^2}{\eta^2} = \psi^2\lambda. \tag{25}$$

While the linear growth of the variance may commonly appear in supply chain literature, it is by no means the only relation found in many *physical* processes. Indeed, a free "quantum" particle (i.e., a particle with zero potential) does *not* exhibit quadratic variation. This is significant because quantum physics is a branch of physics in which randomness is intrinsic and fundamental. Instead, the variance of the location of a free quantum particle increases as t^2.

6. A General Model for Production-Inventory Systems

Equation (17) demonstrates how a force can create a net-flow that is demand only. This example represents an uncontrolled system in which the expectation of the net-flow does not change. The potential to generate such a force is $V(q) = F_0 q = \eta\lambda q$. We can also create state-dependent potentials that can be used to control inventory.

6.1. The General Equation

Using such potentials, the general equations for a PI system can be derived using Hamilton's equations combined with the elements of a Brownian motion.

DEFINITION. General equations of a production-inventory system.

The following equations describe the behavior of a general production-inventory system with a Hamiltonian of $\mathcal{H} = p^2(t)/2 + V(q)$, a damping coefficient of η, and a diffusion coefficient of β and where $F(q) = -\partial\mathcal{H}/\partial q$ yielding,

$$\begin{aligned} dp &= (F(q) - \eta p)dt + \beta dW \\ dq &= pdt \\ p(0) &= p_0 \\ q(0) &= q_0 \end{aligned} \tag{26}$$

With this we can immediately write the steady-state distribution of the net-flow and net-inventory, as

$$\rho(q,p) = C\exp\left(-\frac{\mathcal{H}(q,p)}{kT}\right), \tag{27}$$

where

$$C = \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\rho(q,p)dqdp\right)^{-1}$$

and $kT = \beta^2/(2\eta)$. Equations (26) can be used to determine the transient behavior of such systems, whereas (27) can be used to determine probabilities along with the expected values and variances of p and q. For more complex potentials, the equations of (26) provide an easy means to simulate such a system (see, e.g., Burrage et al. 2007 for an excellent treatment of the numerical integration of SDE's).

6.2. Example: Damped Harmonic Oscillator

The harmonic oscillator is a classic physics problem. The Hamiltonian is $\mathcal{H} = p^2/2 + \omega^2q^2/2$. In our setting, the corrective force is proportional to the difference between the net-inventory and its target. The quadratic potential and linear force in this type of control is reminiscent to the control proposed in the classic study by Holt et al. (1960). The previous example involved two first-order differential equations, but could have been written as a single second-order differential equation in terms of q, as $\ddot{q} + \eta\dot{q} = F(q) + B(t)$, where $F(q)$ is determined by the potential and $B(t)$ is an external Brownian motion force. A simpler differential equation with a corrective force that is proportional to the displacement from a target (here, without loss of generality, set to zero) is given by

$$\ddot{q} + \eta\dot{q} + \omega^2 q = \beta W(t). \tag{28}$$

Such an ordinary second-order differential equation with constant coefficients describes the motion of a "damped harmonic oscillator (DHO)," where ω is a "natural frequency" and a driving force, $\beta W(t)$. Note that the restoring force is $-\omega^2 q$.

The behavior of the system depends greatly on the natural frequency of the system. If $\omega < \eta/2$, the system will be "over-damped" and will take longer to reach equilibrium. On the other hand, if $\omega > \eta/2$, the system is said to be "under-damped" and will oscillate above and below the target before converging. However, if $\omega = \eta/2$, the system is "critically damped" and will reach equilibrium in minimum time without going below zero. See the online Appendix for a complete derivation of the solutions and the steady-state results along with numerous examples. However, we need not integrate because we can obtain the steady-state variances directly from the Hamiltonian,

$$\begin{aligned} \rho(q,p) &= C\exp\left(-\frac{\mathcal{H}(q,p)}{kT}\right) \\ &= \frac{1}{2\pi kT}\exp\left(-\frac{p^2+\omega^2q^2}{2kT}\right), \end{aligned}$$

and identifying $kT = \langle p^2\rangle = \beta^2/(2\eta)$. This result indicates that the steady-state mean values of both q and p are zero and their variances are

$$\begin{aligned} \lim_{t\to\infty}\mathrm{var}(q(t)) &= \frac{\beta^2}{2\eta\omega^2} = \frac{kT}{\omega^2} \\ \lim_{t\to\infty}\mathrm{var}(p^2(t)) &= -\frac{\beta^2}{4a} = \frac{\beta^2}{2\eta} = kT. \end{aligned}$$

7. Implementing the Hamiltonian Control

The above analysis assumes that the system is controlled continuously. This would mean constantly adjusting production in an attempt to keep the net-inventory on target. Of course, this would be impractical in a real production-inventory system where changes are made in discrete time increments.

However, if we naively attempt to implement such a strategy, the system could become unstable as well as unpredictable. Fortunately, this problem is not unlike that of numerically integrating a SDE where the issues of stability and predictability are paramount. Thus, using such numerical methods to periodically recalculate production rates should be both practical and predictable.

7.1. Control Methods from Numerical Integration of SDE

An easy method is the "modified leap-frog" method by Mannella (2006) that can be implemented in our notation as,

$$\begin{aligned}\hat{Q} &= Q_n + \frac{1}{2}P_n\Delta t\\ P_{n+1} &= c_2\Big(c_1P_n + f(\hat{Q})\Delta t + \beta\Delta W_n\Big)\\ Q_{n+1} &= \hat{Q} + \frac{1}{2}P_{n+1}\Delta t,\end{aligned}$$

where $c_1 = 1 - \frac{1}{2}\eta\Delta t$, $c_2 = \left(1 + \frac{1}{2}\eta\Delta t\right)^{-1}$, and $f(\hat{Q}) = -\frac{\partial V(q)}{\partial q}\Big|_{\hat{Q}}$. Note that W_n is a normal random variable with zero mean and variance of Δt. Applying this method to the DHO results in the correct steady-state variance of q while the error in var(p) is independent of η.

The harmonic oscillator is important because any system in equilibrium will behave, as a first-order approximation, like a harmonic oscillator. Thus, the behavior of a harmonic oscillator in equilibrium will be a reasonable approximation of the behavior of almost any potential for small perturbations. So, let $f(q) = -\omega^2 q$ and define the average demand, D_n, and production, X_n, as

$$D_n = c_2(\eta\theta\Delta t - \beta W_n) \tag{29}$$

$$X_n = c_2\Big(-\omega^2\hat{Q} + \eta\theta\Big)\Delta t, \tag{30}$$

where θ and β will be selected to match the demand parameters. Note that D_n contains all the randomness and X_n depends only on previous performance. Then,

$$P_{n+1} = c_1c_2P_n + X_n - D_n. \tag{31}$$

The average demand determines θ

$$\begin{aligned}\langle D\rangle = \lambda &= c_2\eta\theta\Delta t\\ \lambda &= \frac{\eta\theta\Delta t}{1 + \frac{1}{2}\eta\Delta t,}\end{aligned}$$

so that

$$\theta = \frac{\lambda}{\eta\Delta t}\left(1 + \frac{1}{2}\eta\Delta t\right). \tag{32}$$

Likewise,

$$\begin{aligned}\mathrm{var}(D\Delta t) = c_2^2\beta^2(\Delta t)^3 &= \psi^2\lambda\Delta t\\ \frac{\beta^2(\Delta t)^2}{\left(1 + \frac{1}{2}\eta\Delta t\right)^2} &= \psi^2\lambda\end{aligned}$$

or

$$\beta^2 = \frac{\psi^2\lambda\left(1 + \frac{1}{2}\eta\Delta t\right)^2}{(\Delta t)^2}. \tag{33}$$

7.2. Variances in the Discrete Time Model

The above method is guaranteed to yield the same variance for Q as the theoretical continuous case

$$\mathrm{var}(Q) = \frac{\beta^2}{2\eta\omega^2} = \frac{kT}{\omega^2}. \tag{34}$$

The variance of the production rate, X, is

$$\begin{aligned}\mathrm{var}(X) &= \mathrm{var}\Big(c_2(-\omega^2\hat{Q})\Delta t\Big)\\ &= c_2^2\omega^4(\Delta t)^2\mathrm{var}(\hat{Q})\\ &= c_2^2\omega^4(\Delta t)^2\mathrm{var}(Q + \frac{1}{2}P\Delta t)\\ &= c_2^2\omega^2(\Delta t)^2 kT\left(1 + \frac{\omega^2(\Delta t)^2}{4}\right).\end{aligned} \tag{35}$$

Then, the uncertainty relation will be

$$\sigma_X\sigma_Q = c_2kT\left(1 + \frac{\omega^2(\Delta t)^2}{4}\right)^{1/2}\Delta t. \tag{36}$$

7.3. Relation Between the Hamiltonian and the Base Stock Model

It would appear that the Hamiltonian formulation of the DHO is completely different from our original BSM. The BSM is a first-order SDE, whereas the DHO is a second-order one. However, after converting into discrete time, we find that the BSM is a special case of the Hamiltonian model.

Note that if $\eta\Delta t = 2$, then $c_1 = 0$ and $c_2 = 1/2$ and

$$\theta = \frac{\lambda}{2}(1 + 1) = \lambda. \tag{37}$$

Likewise,

$$\begin{aligned}\mathrm{var}(D\Delta t) = c_2^2\beta^2(\Delta t)^3 &= \psi^2\lambda\Delta t\\ \frac{\beta^2(\Delta t)^2}{2^2} &= \psi^2\lambda\\ \frac{\beta^2}{\eta^2} &= \psi^2\lambda\end{aligned}$$

or

$$\begin{aligned}\beta^2 &= \frac{\psi^2\lambda\left(1 + \frac{1}{2}\eta\Delta t\right)^2}{(\Delta t)^2}\\ &= \frac{\psi^2\lambda(2)^2}{(\Delta t)^2}\\ &= \psi^2\lambda\eta^2.\end{aligned}$$

Moreover, it is easy to show that the equation for $\hat{Q}$ is *exactly* the equation for the BSM.

$$
\begin{aligned}
\hat{Q}_n &= Q_n + \frac{1}{2}P_n\Delta t \\
P_{n+1} &= \frac{1}{2}\left(-\omega^2\hat{Q}_n\Delta t + \beta\sqrt{\Delta t}Z\right) \\
Q_{n+1} &= \hat{Q}_n + \frac{1}{2}P_{n+1}\Delta t \\
\hat{Q}_{n+1} &= Q_{n+1} + \frac{1}{2}P_{n+1}\Delta t \\
&= \hat{Q}_n + P_{n+1}\Delta t \\
\hat{Q}_{n+1} &= \hat{Q}_n + \frac{1}{2}\left(-\omega^2\hat{Q}_n\Delta t + \beta\sqrt{\Delta t}Z_n\right)\Delta t \\
&= \hat{Q}_n - g\hat{Q}_n + \psi\sqrt{\lambda\Delta t}Z_n
\end{aligned}
\tag{38}
$$

where and

$$
\begin{aligned}
g &= w^2\Delta t/2 \\
\frac{\beta^2(\Delta t)^2}{4} &= \frac{\beta^2}{\eta^2} = \psi^2\lambda
\end{aligned}
$$

Since $\hat{Q}_n = Q_n + P_n\Delta t/2$ and since, in equilibrium, Q and P are independent, normally distributed random variables, the variance of $\hat{Q}$ will always be greater than the variance of Q. Consequently, the BSM will always have greater variance than that for the equivalent DHO.

7.4. Comparison of the Variances of the Discrete Systems

When the two models are equivalent, the variance of net-inventory for the discrete time DHO can be written in terms of $\psi^2\lambda$ and g, by noting $\eta = 2/\Delta t$ and $\omega^2 = 2g/\Delta t$, viz.,

$$
\mathrm{var}(Q) = \frac{\beta^2}{2\eta\omega^2} = \frac{\psi^2\lambda}{\omega^2\Delta t} = \frac{\psi^2\lambda}{2g}.
$$

We can use these relations and (35) to obtain the variance of production.

$$
\mathrm{var}(X) = \frac{\psi^2\lambda}{2}g\left(1 + g\frac{\Delta t}{2}\right) \tag{39}
$$

so that,

$$
\sigma_X\sigma_Q = \frac{\psi^2\lambda}{2}\left(1 + g\frac{\Delta t}{2}\right)^{1/2}. \tag{40}
$$

We can also obtain the variances of the BSM from

$$
\begin{aligned}
\mathrm{var}(Q_b) &= \mathrm{var}(\hat{Q}) \\
&= \frac{\psi^2\lambda}{2}\frac{\eta}{\omega^2}\left(1 + \frac{\omega^2(\Delta t)^2}{4}\right) \\
&= \frac{\psi^2\lambda}{2g}\left(1 + g\frac{\Delta t}{2}\right).
\end{aligned}
$$

Likewise, the variance of X_b,

$$
\begin{aligned}
\mathrm{var}(X_b) &= g^2\mathrm{var}(Q_b) \\
&= \frac{\psi^2\lambda}{2}g\left(1 + g\frac{\Delta t}{2}\right).
\end{aligned}
$$

Thus,

$$
\sigma_{X_b}\sigma_{Q_b} = \frac{\psi^2\lambda}{2}\left(1 + g\frac{\Delta t}{2}\right). \tag{41}
$$

Note that, for positive values of g, (40) will always be less than (41).

7.5. Example Simulation of Control

Now suppose we have Poisson demand with an average of 16 units per period so that $\psi^2 = 1$ and $\lambda = 16/$day and we use the basic control with $g = 1$. To match the traditional model we set $\eta\Delta t = 2$. This implies $\theta = \lambda = 16/\mathrm{day}$. For stability we choose $\Delta t = 1/3$, so that $\eta = 6$. Then, for the DHO,

$$
\begin{aligned}
\mathrm{var}(Q_D) &= \frac{\psi^2\lambda}{2g} = 8.000 \\
\mathrm{var}(X_D) &= g\frac{\psi^2\lambda}{2}\left(1 + g\frac{\Delta t}{2}\right) = 9.333,
\end{aligned}
$$

so that,

$$
\sigma_{X_D}\sigma_{Q_D} = 8.641.
$$

And for the BSM,

$$
\begin{aligned}
\mathrm{var}(Q_B) &= \frac{\psi^2\lambda}{2g}\left(1 + g\frac{\Delta t}{2}\right) = 9.333 \\
\mathrm{var}(X_B) &= g\frac{\psi^2\lambda}{2}\left(1 + g\frac{\Delta t}{2}\right) = 9.333
\end{aligned}
$$

and

$$
\sigma_{Q_B}\sigma_{X_B} = 9.333.
$$

The simulation involves 200,000 replicates over 18 periods of 1/3 each. Since the system begins in steady state (i.e., $Q(0) = 0$, $P(0) = 0$), there is no transient period. The results shown at the top of Table 1 confirm the fact that the Mannella algorithm converges to the correct value of the variance of Q and is very close to the variance of X as well. Regardless, the simulation is actually the "correct" value when we are controlling a real system at points in time (as opposed to a continuous control). Thus, the BSM has roughly 8%–9% more variability than the damped harmonic oscillator model (comparing the products of the simulated standard deviations).

Table 1 Simulation Results for the Base Stock Model (BSM) and Equivalent Damped Harmonic Oscillator (DHO)

Measure	Simulation	Theoretical	% Error
Matched DHO and BSM			
DHO var(Q)	7.999	8.000	−0.02
DHO var(X)	9.628	9.333	3.15
BSM var(Q)	9.628	9.333	3.15
BSM var(X)	9.578	9.333	2.63
DHO $\sigma_Q\sigma_X$	8.775	8.641	1.56
BSM $\sigma_Q\sigma_X$	9.603	9.333	2.89
$\sigma_Q\sigma_X$ difference	8.01%	9.43%	
Enhanced DHO and BSM			
DHO var(Q)	8.536	8.533	0.04
DHO var(X)	5.774	5.600	3.11
BSM var(Q)	9.628	9.333	3.15
BSM var(X)	9.578	9.333	2.63
DHO $\sigma_Q\sigma_X$	7.021	6.913	1.56
BSM $\sigma_Q\sigma_X$	9.603	9.333	2.89
$\sigma_Q\sigma_X$ difference	36.78%	35.02%	

Can we do better? It appears so. If we set $\eta = 10$ and run the DHO model. The variance of both measures is predicted to drop to 8.5 for Q and 5.6 for X, yielding a product of standard deviations of 6.9 as compared to 9.3 for the BSM. The simulation generated a value of 7.0 for the same product. Thus, the improvement over the base stock model is around 35%.

8. Conclusions

We have seen that the Hamiltonian formulation of the Least Action model can be applied to the supply chain with interesting results. It appears that the traditional method of controlling production to hit an inventory target is a special case of this Hamiltonian model. As such, the Hamiltonian model cannot perform worse and, indeed, can perform as a more efficient production control method.

Adding randomness to the Hamiltonian model quite "naturally" results in a Brownian motion. Such models have seen a great deal of application in the supply chain, but were typically postulated as a reasonable a priori model and not derived from basic principles. Our development here provides additional justification for such models.

This exploration has also suggested new ways to control production and inventory. Whereas most of previous inventory models have been limited in the number of possible probability distributions, the possibility of using a variety of "potential" functions as controls suggests an almost limitless number of ways to provide control. Moreover, the distribution of net-inventory of any such control is easily constructed from the potential used,

$$\rho(q) = C\exp\left(-\frac{2\eta V(q)}{\beta^2}\right),$$

where C is used to normalize the distribution.

Finally, the proposed model links the three "buffers" of inventory, time, and capacity in a "canonical" way. Although the link between the time and inventory buffer was already well understood (while backorder time is proportional to negative net-inventory, see e.g., Zipkin 2000), the link between these two and capacity has been studied heretofore only for special cases. We have now seen a direct link between production and net-inventory as these two are "canonical conjugate" variables and have an associated "uncertainty principle" that is directly related to the idea of buffers mitigating variability.

$$\sigma_x\sigma_q \geq \frac{\psi^2\lambda}{2}.$$

8.1. Future Work

We have only begun to scratch the surface of the Least Action model. Other areas of exploration include:

- Optimization of the supply chain: The uncertainty relation shows that if we want tighter control of inventory, we must accommodate a larger variance in production and therefore have a larger capacity "buffer." Conversely, if we cannot afford the extra capacity, we will have to live with either larger inventories or worse customer service. This relation provides a quantitative means to optimize the supply chain considering the cost of capacity, service (time), and inventory in a way to minimize total life cycle costs. This should be explored in future work.
- How is the Hamiltonian model related to system dynamics models such as those developed by Forrester (1961)and more recently explored by, Sterman (2000)? Both the Hamiltonian model and system dynamics involve differential equations and damping. Is the Hamiltonian control simply a coincidental example of a systems dynamics model? If so, is there an underlying structure?
- Conservation and Symmetry: The relation between symmetry and conservation laws runs deep in physics. Noether's theorem (see the online Appendix for a short proof) suggests that for every symmetry in the Hamiltonian there must be some conservation law. An interesting question is whether there are similar symmetries and conservation laws in supply chain dynamics. If not, why not? If so, what are they?

- The nature of randomness in a supply chain: As we mentioned above, the branch of physics that deals with intrinsic randomness is quantum physics. Moreover, quantum physics does not exhibit quadratic variation, but it appears that supply chain processes do. Nonetheless, our derivation was based on "statistical physics" which also exhibits quadratic variation. So the question is whether the supply chain is intrinsically random or simply unknown.
- What is the best potential: to use to control that will minimize variability?
- Other Applications: Can the Least Action Principle be applied to other areas of operations management?

9. Acknowledgments

The author would like to thank Mr. Edward S. Pound of Factory Physics, Inc., and Dr. Antonio Arreola-Risa of Texas A&M University for their willingness to listen to a bunch of esoteric physics ideas. I also want to thank them for their keen insights, thoughtful comments, and discerning suggestions which greatly improved the manuscript. I am extremely grateful to have Professor Wallace Hopp of the University of Michigan as a colleague and friend over these past 25 years. Working with him at Northwestern University during the 1980s and 1990s led to the development of basic factory physics concepts. I would also like to thank Professor Kalyan Singhal for his suggestions in making this study more practical. Thanks also go to William Spearman of Harvard University for his suggestion to consider statistical physics as an alternative to the quantum formulation. Finally, I would like to thank my wife, Blair, for her patient listening to what must have seemed like the ravings of a crazy former physicist.

Notes

[1]This was probably not true 100 years ago and is certainly untrue today. Ironically, Rutherford received his Nobel Prize in chemistry, not physics.

[2]To simplify the notation we will use the same notation as that employed in control theory (and first developed by Newton) indicating a generalized coordinate with q, its time derivative as $\dot{q}$, and its second derivative as $\ddot{q}$. Moreover, we will not explicitly indicate the time dependence of q and its derivatives, as it would make the notation clumsy.

[3]Lagrange multipliers found in the operations research literature are related to the classical Lagrangian. In classical dynamics, similar multipliers are used to constrain motion in much the same way as the OR multipliers are used to enforce constraints in optimization .

[4]Indeed, the approximate model for Brownian motion was used by Jean Perrin to determine the value of Avogadro's number, an accomplishment for which he won the Nobel Prize in 1926.

References

The ATLAS Collaboration. 2012. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC, *Phys. Lett. B* **716**: 1–29.

Boltyanskii, V. G. 1958. Maximum principle in the theory of optimal processes, *Dokl. Akad. Nauk. SSSR* **119**(6): 1070–1073.

Burns, J. B., C. I. Connolly, R. A. Grupen, J. S. Jog. 2012. A least-action model for dyskinesias in Parkenson's disease. Available at http://www.ai.sri.com/pubs/files/388.ps (accessed date November 4, 2012).

Burrage, K., I. Leane, G. Lythe. 2007. Numerical methods for second-order stochastic differential equations, *SIAM J. Sci. Comput*. **29**: 245–264.

Forrester, J. W. 1961. *Industrial Dynamics*. Pegasus Communications, Waltham, MA.

Harrison, J. M. 1985. *Brownian Motion and Stochastic Flow Systems*. Wiley, New York.

Hopp, W. J., M. L. Spearman. 2011. *Factory Physics*, 3rd edn. Waveland Press, Long Grove, IL.

Holt, C. C., F. Modigliani, J. F. Muth, H. A. Simon. 1960. *Planning Production, Inventories, and Work Force*. Prentice-Hall, Englewood Cliffs, NJ.

Kardar, M. 2007. *Statistical Physics of Particles*. Cambridge University Press, New York.

Kaila, V. R. I., A. Annila. 2008. Natural selection for least action, *Proc. R. Soc. A* **464**: 3055–3070.

Langevin, P. 1908. Sur la théorie du mouvement brownien, *Comptes-rendus de l'Académie des Sciences* **146**: 530–532.

Kubo, R. 1966. The Fluctuation-dissipation theorem, *Rep. Prog. Phys.* **29**: 255–284.

Mannella, R. 2006. Numerical stochastic integration for quasi-symplectic flows, *SIAM J. Sci. Comput.* **27**: 2121–2139.

Oksendal, B. 1995. *Stochastic Differential Equations*. Springer-Verlag, Berlin.

De Sapio, V., O. Khatib, S. Delp. 2007. Least action principles and their application to constrained and task-level problems in robotics and biomechanics, *Multibody Syst. Dyn.* **19**: 303–322.

Sethi, S. P., G. L. Thompson. 2000. *Optimal Control Theory: Applications to Management Science and Economics*, 2nd edn. Kluwer, Norwell, MA.

Samuelson, P. A. 1970. Maximum principles in analytical economics, Nobel Memorial Lecture, December 11, 1970, Stockholm.

Samuelson, P. A. 1974. A biological least-action principle for the ecological model of Volterra-Lotka, *Proc. Natl. Acad. Sci. U S A*, **71–8**: 3041–3044.

Sterman, J. 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. McGraw-Hill, New York.

Stevens, C. F. 1995. *The Six Core Theories of Modern Physics*. MIT Press, Cambridge, MA.

Zipkin, P. H. 2000. *Foundations of Inventory Management*. McGraw-Hill, New York.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix http://www.factoryphysics.com/documents/POMS%20Full%20Online%20Version.pdf