Investigating sustained oscillations in nonlinear production and inventory control models

Virginia L. M. Spiegler, Mohamed M. Naim

Abstract

Even in a deterministic setting, nonlinearities can yield unexpected dynamic behaviours in a production and inventory control system, such as sustained oscillations or limit cycles. Describing function in combination with simulation is used to analyse the effect of discontinuous nonlinearities on the system responses. Utilising a nonlinear production and inventory control model, we investigate the occurrence of limit cycles and propose a technique to predict their amplitude, frequency and stability and to control such oscillations. Findings suggest that, even for an autonomous production and inventory control system, limit cycles do occur and this periodic behaviour occurs due to non-negativity constraint in the ordering rule. Moreover, we demonstrate the potential of the describing function method to provide insight into the impact of system constraints and therefore facilitate a more effective system design. This paper fills a gap in the literature on nonlinear supply chain dynamics by expanding and complementing the sparse recent research in this area. Most previous studies have either focused on linear mathematical models or relied on simulation, which greatly limit the relevancy and/or rigour of the published results.

Keywords: Inventory; nonlinear supply chain dynamics; describing function; rogue seasonality; limit cycle

1. Introduction

System dynamics play a critical role in supply chain performance. These dynamics are normally driven by the application of different production and inventory control system policies and are also considered to be a cause of mismatch between customer demand and supply (Towill et al., 1992; Sterman, 2000; Colicchia et al., 2010; Spiegler et al., 2012). Examples of such dynamics in supply chains include the bullwhip effect (amplified form of order information as it propagates upstream) (Lee et al., 1997), rogue seasonality (cyclic...
disturbances endogenously generated) (Thornhill and Naim, 2006; Shukla et al., 2012) and the ripple effect (impact of a disruption as it ripples through the supply chain network) (Ivanov et al., 2014b,a). System dynamics intervention has been promoted to identify how structure and decision policies generate system behaviour and to implement structural and policy-oriented solutions (Forrester, 1961; Saleh et al., 2010).

In this paper we are concerned with detecting and controlling rogue seasonality, more specifically limit cycles, which are sustained oscillations induced by nonlinearities. In practice, rogue seasonality has been evidenced, for instance, in the grocery (Shukla, 2010) and steel (Thornhill and Naim, 2006) industries, but determining the root causes for such oscillations is still a challenge. In the steel industry, Thornhill and Naim (2006) used a data-driven technique to separate rogue seasonality from cyclical disturbances externally generated by the customer demand. Although they were able to establish that these oscillations were generated by the internal production planning systems, their technique was unable to find the root causes. Capital intensive industries, such as steel and chemical, are characterised by several constraints in their manufacturing process (Thornhill and Naim, 2006), therefore, these complex structures with nonlinearities could potentially induce oscillations.

The aim of this paper is therefore to propose a methodological framework to investigate the occurrence of limit cycles, characterised by sustained and periodic oscillations, induced by nonlinearities in production and inventory control systems, such as limits to capacity or non-negativities, and to devise a suitable compensation method. As propagated through the supply chain, oscillations have been associated with increased costs and poor service levels (Towill and Vecchio, 1994; Disney et al., 2004; Villegas and Smith, 2006) and therefore need to be properly understood and controlled.

2. Literature review

In the supply chain literature, a plethora of studies have made efforts to understand system dynamics and design production and inventory control systems that reduce demand amplification (Chen et al., 2000; Dejonckheere et al., 2003; Disney and Towill, 2003; Chen and Lee, 2012; Wang et al., 2015b) and decrease risk of stock-out (Wang et al., 2015a; Spiegler et al., 2016b), consequently improving transport and production operations (Bicheno et al., 2001; Potter and Lalwani, 2008; Juntunen and Juga, 2009; Cannella and Ciancimino, 2010;
Spiegler and Naim, 2014) and service and financial performance (Torres and Maltz, 2010; Ivanov and Sokolov, 2013; Dai et al., 2016).

Among the methodological approaches used to analyse and design production and inventory control systems are statistical inventory control, system dynamics simulation and control theory (Dejonckheere et al., 2003; Towill et al., 2007). The former approach, although requiring complex mathematics, is concerned with understanding signal statistics of relatively simple systems in order to find optimum performance solutions for supply chain costs and service levels. For instance, Hosoda and Disney (2006) analysed a two-echelon supply chain inventory problem and showed that individual optimum ordering policies do not necessarily yield globally optimal solutions for the overall supply chain. Another example is Dai et al. (2016) who looked at different optimum ordering decisions made by the retailer and manufacturer when given information with different levels of quality regarding inventory inaccuracies.

In contrast, system dynamics simulation enabled a complex model to be built from causal loop diagrams and cause-and-effect relationships to be identified by comparing various graphical outputs of various scenarios. These simulation techniques have been advocated to investigate complex, high-order, nonlinear supply chain models (Forrester, 1961; Sterman, 2000). Notable and recent applications of system dynamics simulation to solve production and inventory control problems include the works of Goncalves et al. (2005); Ali and Deif (2016). The former make use of simulation to analyse the complex hybrid push-pull production system of Intel, while Ali and Deif (2016) proposes a novel integrated metric to measure the degree of leanness of a dynamic manufacturing model. On the other hand, when analysing simpler models linear control theory is recommended to capture the dynamics of the system’s responses through analytical tools, such as signal flow diagrams, block diagram manipulation and simplification, Laplace/z-transforms and transfer function analysis, “hard system” control laws and frequency response plots. Examples of linear control theory application in production and inventory control can be found in Towill (1982); John et al. (1994); Disney et al. (2004); Zhou et al. (2010). Despite providing a more rigorous and comprehensive understanding of the system’s behaviour, linear control theory has been criticised for its limited applicability (Spiegler et al., 2012; Wang et al., 2014) since it disregards the existence of nonlinearities. In supply chains, nonlinearities can naturally occur through the existence
of physical and economic constraints (Spiegler et al., 2016a) and should not be disregarded due to the unexpected behaviour they may generate in the system (Forrester, 1961).

In order to overcome the exhausting and challenging task of simulating complex models without preliminary analysis and to recognise the effects of nonlinear dynamics previously neglected by traditional control theorists, recent articles have demonstrated the benefit of applying modern or nonlinear control theory for production and inventory control design (Wang and Disney, 2012; Wang et al., 2014, 2015b; Spiegler et al., 2016a,b). By conducting a thorough literature review on nonlinear system dynamics, Spiegler et al. (2016a) suggested appropriate nonlinear control theory methods for different types of nonlinearities, such as continuous or discontinuous, single- or multi-valued. They explored different nonlinear control structures by using the often cited Forrester production-distribution model as a benchmark supply chain system and provided a systematic method for the rigorous analysis and design of nonlinear supply chain dynamics models. Wang and Disney (2012) and Wang et al. (2014) used graphical and eigenvalue methods to explore stability boundaries of a piecewise linear inventory control system and to identify a set of behaviours in the unstable region, such as periodicity and chaos. When trying to understand nonlinear behaviour within the system’s stable region, Wang et al. (2015b) applied the describing function technique to analyse the bullwhip effect under non-negative ordering policies. Similarly, Spiegler et al. (2016b) used the describing function method to analyse the resilience performance of a grocery supply chain under non-negative batching constraints in the ordering rule and backlog situations in the shipment estimation. In both the works of Wang et al. (2015b) and Spiegler et al. (2016b), the describing function method enabled understanding of the nonlinear impact of different demand signals, with different means, amplitudes and frequencies, on supply chain performance.

In this paper, we expand on the works of Wang et al. (2015b) and Spiegler et al. (2016b) and follow the recommendations of Spiegler et al. (2016a) on the application of describing function method to not only understand the impact of different demand signals passing through replenishment constraints but also to use it as a design tool to control and avoid sustained oscillations or limit cycles that are intrinsic to some nonlinear systems and, therefore, managing rogue seasonality. We also complement the works of Wang and Disney (2012) and Wang et al. (2014) whose methods, despite enabling the identification of oscillatory re-
gions in the production and inventory control system, did not reveal the characteristics of
the limit cycling behaviour. “Although precise knowledge of the waveform of a limit cycle is
usually not mandatory, knowledge of the existence of a limit cycle, as well as its approximate
amplitude and frequency, is a prerequisite to good system design” (Gelb and Velde, 1968).

3. Model formulation: Nonlinear production and inventory control model

The order-up-to (OUT) policy is a popular production and inventory control model. By
reviewing the system’s inventory position every period, an order is placed to bring the inven-
tory position ‘up-to’ a target level. We chose to analyse the Automatic Pipeline Inventory
and Order Based Production Control System (APIOBPCS) (John et al., 1994), which is
very similar to the classical discrete OUT system (Disney and Towill, 2002). This system
takes into account a demand forecasting method, production and distribution lead-times,
an inventory feedback loop, a work in process (WIP) feedback loop and target inventory
levels. This model and its variants have been well-acknowledged in the production control
and supply chain literature (Cannella et al., 2008; Aggelogiannaki et al., 2008; Cannella and
Ciancimino, 2010; Eshlaghy and Razavi, 2011; Georgiadis and Michaloudis, 2012).

Figure 1 illustrates the nonlinear block diagram representation of the APIOBPCS. Previous
authors have used nonlinear forms of the APIOBPCS in order to make this model more
representative of a real production and inventory control system (Shukla et al., 2009; Wang
and Disney, 2012; Spiegler et al., 2012; Wang et al., 2014). By inserting CLIP functions to
avoid order rates reaching negative values (also referred to as forbidden return constraint)
and shipments being made without any on-hand inventory, the model in Figure 1 has been
used to replicate the Beer Game (Spiegler and Naim, 2014). Other OUT systems that con-
sider some of these nonlinearities include Laugesen and Mosekilde (2006); Mosekilde and
Laugesen (2007); Wang and Disney (2012); Hwarng and Yuan (2014); Wang et al. (2015b).

More recently, a similar model has been identified in the replenishment system of a grocery
retailer (Spiegler et al., 2016b), but, in addition to the non-negativity constraint, orders were
also batched based on the buying quantities and truckload limits.

The production and inventory control model in Figure 1 is characterised by three control
parameters $T_d$, $T_i$, and $T_w$ that drive the demand, inventory and work in process (WIP)
policies, respectively, and a physical parameter, the actual lead-time $T_p$. The expected lead-
Figure 1: Block diagram of the production and inventory control model

time $\bar{T}_p$ is assumed to be equal to the actual lead-time. In the demand policy, the value of the current demand $d(t)$ is exponentially smoothed. Hence, the parameter $T_a$ represents the time to average demand so that the exponential smoothing constant is $\alpha = 1/(1 + T_a/\Delta t)$, where $\Delta t$ is the sample time interval.

The inventory and WIP policies are characterised by feedback loops (a feedback of information that compares current values of inventory and WIP with their target values). The inventory control is concerned with the rate $(1/T_i)$ at which a deficit in inventory is recovered. This policy is responsible for reducing the discrepancy between desired inventory $d_i$ and actual inventory $i(t)$ or backlog $b(t)$. The pipeline policy considers the actual work in process $w(t)$ and the time $T_w$ it takes to recover to target levels. While the $d_i$ is a constant value, the desired WIP $d_w(t)$ is function of $\bar{T}_p$ and the average demand $\hat{d}(t)$.

Finally, the order $o(t)$ placed onto the supplier will take into account $\hat{d}(t)$, the errors in $i$ and $w$ and the non-negativity constraint. The receipt $r(t)$ of material is represented by a first order lag with a time constant of $T_p$. The shipment sent $s(t)$ to the customer will depend on the desired shipment $d(t) = d(t) + b(t)$ and the maximum shipment $ms(t) = r(t) + i(t)$. 
4. Describing function method

Since the model in Figure 1 contains discontinuous nonlinearities, the describing function method can be used to find the quasi-linear representation of these nonlinearities. The first discontinuity to be analysed is the CLIP function present in the ordering system which prevents orders from assuming negative values. Then, the constraint in the shipments will be investigated. This section follows the systematic method proposed by Spiegler et al. (2016a).

4.1. Non-negativity constraint

Considering the nonlinearity in the ordering system with an input \( do(t) \) and an output \( o(t) \). In an open-loop system, the sinusoidal input function

\[
do(t) = A_{do}.\cos(\omega t) + B_{do},
\]

where \( \omega \) is the angular frequency, \( A_{do} \) is the amplitude and \( B_{do} \) is the mean, will produce an output \( o(t) \) of the same \( \omega \) but with different amplitude and mean. The output \( o(t) \) does not depend on the past values of \( do(t) \), but it varies according to the actual state of \( do(t) \). Hence, \( o(t) \) can be represented by the following piecewise linear equation:

\[
o(t) = \begin{cases} 
do(t), & \text{if } -\gamma < \omega t < \gamma \\
0, & \text{if } -\pi < \omega t < -\gamma \text{ and } \gamma < \omega t < \pi,
\end{cases}
\]

where \( \gamma = \cos^{-1}\left(\frac{-B_{do}}{A_{do}}\right) \). Equation 2 is asymmetric and can be approximated to

\[
o(t) \approx N_{A_{do}}.A_{do}\cos(\omega t + \phi) + N_{B_{do}}.B_{do},
\]

where \( \phi, N_{A_{do}} \) and \( N_{B_{do}} \) are the phase shift, the gains in amplitude and mean, respectively. In order to determine \( \phi, N_{A_{do}} \) and \( N_{B_{do}} \), we need to expand the series and determine its harmonic coefficients. The Fourier series expansion method is used to represent the output as a series, such as:

\[
o(t) \approx b_0 + \sum_{k=1}^{\infty} [a_k\cos(k.\omega t) + b_k\sin(k.\omega t)],
\]
where \( a_k, b_k \) and \( b_0 \) are the Fourier coefficients.

For the describing function, only the fundamental harmonic is usually used to approximate the periodic series. This is appropriate for symmetric systems because they contain only odd harmonics; therefore, higher harmonics will be attenuated by the linear dynamics of the system (Vukic et al., 2003). However, in the case of asymmetric nonlinearities, the second harmonic also occurs. For this reason, describing function techniques tend to be less accurate than those applied to symmetrical systems and the complementary use of simulation is recommended (Atherton, 1975). If we approximate the piecewise linear output \( o(t) \) to the first harmonic, we have:

\[
o(t) = b_0 + a_1 \cos(\omega t) + b_1 \sin(\omega t) = b_0 + \sqrt{a_1^2 + b_1^2} \cos(\omega t + \phi),
\]

where \( \phi = \arctan\left(\frac{b_1}{a_1}\right) \).

In this way, by comparing Equations 3 and 5, we can define the describing function gains as:

\[
N_{A_{do}} = \frac{\sqrt{a_1^2 + b_1^2}}{A_{do}},
\]

\[
N_{B_{do}} = \frac{b_0}{B_{do}}.
\]

For single-valued nonlinearities, the coefficient \( b_1 \), the imaginary part of the describing function, will be equal to zero and therefore \( \phi \) will also be zero. By replacing the Fourier coefficients in Equations 6 and 7, we find

\[
N_{A_{do}} = \frac{\gamma - \cos(\gamma) \cdot \sin(\gamma)}{\pi}
\]

\[
N_{B_{do}} = \frac{B_{do} \cdot \pi + A_{do} \cdot (-\gamma + \pi) \cdot \cos(\gamma) + A_{do} \cdot \sin(\gamma)}{B_{do} \cdot \pi}
\]

Figure 2 illustrates how the coefficients of the describing function vary as \( A_{do} \) increases for any \( B_{do} > 0 \). For values of \( A_{do} \) lower than \( B_{do} \), the system behaves as linear and output \( o(t) \) will be equal to the input \( do(t) \) corresponding to \( N_{A_{do}} = 1 \) (Figure 2a). However, when \( A_{do} \)
increases then only a fraction of this rate will actually be ordered corresponding to $N_{A_{do}} < 1$. By inspecting Equation 8, we find that as $A_{do}$ approaches infinity, $N_{A_{do}}$ approaches 0.5. So, the amplitude gain of the describing function can only vary from 0.5 to 1.

![Amplitude gain](image1.png)

![Mean gain](image2.png)

Figure 2: Terms of describing function for the non-negativity constraint

On the other hand, the value of $N_{B_{do}}$ rises as $A_{do}$ increases because the limit value of the order rate is at its minimum (Figure 2b).

4.2. Shipment constraint

The second nonlinearity in the model is the CLIP function in the shipment system that is used to avoid any shipments being made to the store if no inventory is available. Hence, the output $ss(t)$ will be equal to the input $ds(t)$ if $ds(t) < ms(t)$. While in the ordering system the non-negativity constraint was fixed, the shipment constraint is given by the current responses of $ds(t)$ and $ms(t)$, which vary over time. Because of that, this nonlinearity is not only amplitude-dependent but also frequency-dependent. Therefore, there will be one describing function for each frequency. This complex behaviour makes this nonlinearity multi-valued, therefore the output $ss(t)$ will experience not only a change in amplitude and mean, but also a shift in $\phi_{ds}$.

Matlab$^TM$ and Simulink$^TM$ have been used to automate calculations and find the describing function gains for a set of amplitudes and frequency. Figure 3 illustrates the input frequencies and amplitudes for which the system behaves in a linear fashion, which means that when $N_{A_{ds}} = 1$, $\phi_{ds} = 0$. The dark shaded areas indicate that $N_{A_{ds}}$ is below 1 and thus
the system behaves nonlinearly, which is when frequencies are medium-low and amplitudes are high.

![Graph showing terms of describing function for the shipment constraint](image)

**Figure 3:** Terms of describing function for the shipment constraint

5. Limit cycle analysis

One of the effects that nonlinearities can cause is oscillatory behaviour. For an autonomous system, when the state vector returns to one of its previous values, it must necessarily repeat this motion and so the response will keep recurring indefinitely without reaching a steady state. These oscillations, also known as limit cycles, may occur as the output of the nonlinearity switches from one state to another. In order to guarantee a robust system design, these oscillations should be adequately compensated. In this section, the presence or absence of limit cycles in the production and inventory control model due to the presence of nonlinearities is investigated with the help of the describing function method. When limit cycle is predicted, we also estimate its amplitude, frequency and stability so that a suitable compensation method can be suggested.

5.1. General formulation

Here we present the general formulation for limit cycle analysis as in Gelb and Velde (1968); Cook (1986) and Towill (1970). The general block diagram for a feedback control system with a single nonlinearity is shown in Figure 4.

An external reference input $r(t)$ is compared to a signal $y(t)$ and fed into a linear element with transfer function $G(s)$, generating the input to the nonlinearity $u(t)$. The output of the nonlinearity $f(t)$ is then transferred through a feedback transfer function $H(s)$ back into the
loop with a negative feedback. The nonlinear element can be approximated to the describing function amplitude and mean gains $N_A$ and $N_B$, respectively.

Consider that the reference signal $r$ is constant, so that the system is autonomous. It is assumed that if a limit cycle occurs, it can be adequately approximated by a sinusoidal oscillation, such as

$$u(t) = A \cos(\omega t) + B,$$  \hspace{0.5cm} (10)

and correspondingly

$$f(t) \approx N_A A \cos(\omega t + \phi) + N_B B$$

$$\approx \operatorname{Re}\{N_A\} A \sin(\omega t) + \operatorname{Im}\{N_A\} A \cos(\omega t) + N_B B. \hspace{0.5cm} (11)$$

Hence,

$$y(t) \approx \operatorname{Re}\{H(j\omega)N_A\} A \sin(\omega t) + \operatorname{Im}\{H(j\omega)N_A\} A \cos(\omega t) + H(0)N_B B \hspace{0.5cm} (12)$$

so that, setting

$$u(t) = (r - y(t))G(j\omega), \hspace{0.5cm} (13)$$

when equating coefficients of corresponding terms, we get

$$B + G(0)H(0)N_B B = r \hspace{0.5cm} (14)$$

$$A + \operatorname{Re}\{G(j\omega)H(j\omega)N_A\} A + \operatorname{Im}\{G(j\omega)H(j\omega)N_A\} A = 0. \hspace{0.5cm} (15)$$
Limit cycle will occur if both equations are satisfied. Equation 14 shows the constant term and is particularly relevant for asymmetric nonlinearities. Equation 15 shows the oscillatory term. In this way, permanent oscillation will occur in the system for an $A \neq 0$ and when

$$1 + \text{Re}\{G(j\omega).H(j\omega).N_A\} + \text{Im}\{G(j\omega).H(j\omega).N_A\} = 0$$

$$1 + G(j\omega).H(j\omega)N_A = 0$$

$$G(j\omega).H(j\omega)N_A = -1$$

(16)

Hence, the existence of a limit cycle is then predicted if there is an intersection between the loci, in the complex plane, of $G(j\omega).H(j\omega)$ and $-1/N_A$. In other words, oscillations may occur if the system’s open-loop equation intercepts point -1 on the real axis of the s-plane.

5.2. Predicting occurrence of the limit cycle in the nonlinear APIOBPCS

In order to predict the occurrence of limit cycle we need to firstly obtain the system’s output transfer functions. By replacing the CLIP functions of Figure 1 with the amplitude gains $N_{Ado}$ and $N_{Ads}$ in the ordering and shipment rules, respectively, we can easily determine the following closed-loop transfer functions:

$$\frac{O(s)}{D(s)} = \frac{N_{Ado}(1 + sT_p)(sT_iT_p + T_w + s(T_a + T_i)T_w)}{(1 + sT_a)(sT_i(1 + sT_p)T_w + N_{Ado}(sT_iT_p + T_w))},$$

(17)

$$\frac{I(s)}{D(s)} = \frac{-N_{Ads}T_iT_w(1 + s)(1 + sT_a)(1 + sT_p) + N_{Ado}(-sT_iT_p(N_{Ado} - 1) + N_{Ado}(T_i - sT_a - 1))}{(N_{Ado} + s)(1 + sT_a)(sT_i(1 + sT_p)T_w + N_{Ado}(sT_iT_p + T_w))},$$

(18)

$$\frac{W(s)}{D(s)} = \frac{N_{Ado}T_p(sT_iT_p + T_w + s(T_a + T_i)T_w)}{(1 + sT_a)(sT_i(1 + sT_p)T_w + N_{Ado}(sT_iT_p + T_w))}.$$

(19)

Since $N_{Ads}$ only affects $i(t)$, we will use the transfer function in Equation 18 to predict the occurrence of limit cycles that are possibly caused by both nonlinearities. The open-loop transfer function for $i(t)$ will be designated as $\bar{I}(s)$ and can be found by transforming the closed-loop transfer function of Equation 18 into an open-loop form, resulting in:

12
$T(s) = \left(-N_{Ad}(1 + s)(1 + sT_a)T_i(1 + sT_p)T_w + N_{Ad}(-s(-1 + N_{Ad}(1 + sT_a)T_iT_p + (1 + s(T_a + T_i) + N_{Ad}(T_i - 1 - sT_a))T_w) \right) \left((N_{Ad} + 2N_{Ad}s + s^2)(1 + sT_a)T_iT_w \right.$

\[= (1 + sT_p) + N_{Ad}(s(-1 + s + s^2T_a + N_{Ad}(2 + T_a + 2sT_a))T_iT_p + ((-1 + 2N_{Ad} + s)
\left(1 + sT_a) - (N_{Ad} + s)T_iT_w) \right). \]

By replacing $s = j\omega$ in Equation 20 and setting it equal to -1, the solution shows that there may be limit cycles only in the ordering system and this may occur when:

$$T_w = -N_{Ad}T_p.$$  \hspace{1cm} (21)

Since the values of $N_{Ad}$ vary from 0.5 to 1, sustained oscillations may occur when

$$-T_p \leq T_w \leq -0.5T_p.$$  \hspace{1cm} (22)

Conventionally, the WIP feedback controller ($T_w$) assumes a positive value range. Nonetheless, the impact of negative $T_w$ has been investigated in case of an uninformed and irrational replenishment rule design (Wang and Disney, 2012; Wang et al., 2014). Moreover, negative values of $T_w$ have been previously supported for achieving aperiodic responses in a linear OUT policy system and therefore improving supply chain performance (Disney, 2008). According to Disney (2008), responses are highly damped for $T_w < -T_p$ but changes in the lead-time could potentially make the system operate within the region described by Equation 22. This, in combination with a nonlinear ordering rule, may lead the system to behave in a contradictory manner, that is, undamped and oscillatory.

Note that the nonlinearity in the shipment system, represented by the describing function gain $N_{Ad}$, does not seem to cause any limit cycles. For this reason, the following analysis will only be focused on the nonlinearity in the ordering system.

We still need to investigate the other condition for limit cycles as given by Equation 14. Firstly, we need to manipulate Equation 17 in order to obtain a more familiar form and separate the system’s linear frequency-dependent elements from the amplitude-dependent
describing function. This results in:

\[
O(s) = \frac{N_{Ado}(\frac{1}{1+s.T_a} + \frac{1}{s.T_i} + \frac{T_p}{(1+s.T_a)T_w})}{1 + N_{Ado}(\frac{1}{s.T_i(1+s.T_p)} + \frac{1}{s.T_w} - \frac{T_w}{s.T_w(1+s.T_p)})},
\]

which can then be related to the general form of a feedback transfer function \(\frac{N_{Ado}G(s)}{1+N_{Ado}G(s)H(s)}\), where

\[
G(s) = \frac{T_w + s(T_iT_p + T_aT_w + T_iT_w)}{s^2T_aT_iT_w + sT_iT_w},
\]

\[
H(s) = \frac{(1 + sT_a)(sT_iT_p + T_w)}{(1 + s.T_p)(sT_iT_p + sT_aT_w + sT_iT_w + T_w)}.
\]

Hence, the original APIOBPCS in Figure 1 can be re-arranged as the general block diagram of Figure 4, resulting in Figure 5.

![Figure 5: Simplified block diagram for the ordering rule in the nonlinear APIOBPCS](image)

Now with the block diagram defined in Figure 5, we are able to determine whether the condition given by Equation 14 is also satisfied since this nonlinearity is asymmetric \((B_{do} \neq 0)\). Following the guidelines in Gelb and Velde (1968, pg. 343), considering that \(G(s)H(s)\) contains open-loop integrators (poles at the origin), we find that another condition for limit cycling behaviour is that of \(d = 0\).

5.3. Predicting amplitude, frequency and stability of limit cycles in the ordering system

In the frequency domain, \(G(j\omega)H(j\omega)\) is a complex function of \(\omega\) and \(N_{Ado}\) is a function of amplitude \(A_{do}\), as in Equation 8, the solution of the \(G(j\omega)H(j\omega)N_{Ado} = -1\) gives both the frequency and amplitude of the oscillation. We find that the frequency of the
oscillation is therefore predicted to be:

$$\omega_{do} = \frac{T_w^2}{T_p \sqrt{T_i T_w}}.$$  \hspace{1cm} (26)

For a fixed $T_w$ in relation to $T_p$, the frequency of oscillation decreases as lead-time lengthens and as the parameter $T_i$ increases.

Equation 8 and Figure 2 can be used to find $A_{do}$. When $T_w = -T_p$, the describing function gain $N_{A_{do}} \approx 1$ and the amplitude will be asymptotically equal to the asymmetric bias $B_{do}$, which in turn is proportional to the mean of the reference signal $d$. When $T_w = -N_{A_{do}} T_p$ for $0.5 < N_{A_{do}} < 1$, the corresponding amplitude can be determined by Equation 8, or easily seen in Figure 2. The amplitude of the oscillation increases as $T_w$ changes from $-T_p$ to $-0.5T_p$.

Figure 6 illustrates the Nyquist plots for $G(s).H(s)$ and $N_{A_{do}}$ with different values of the control parameter $T_w$ in relation to a fixed lead-time $T_p = 3$. The control parameters $T_a$ and $T_i$ are set as 6 and 3, respectively. Figure 6a illustrates a stable system since there is no encirclement of the Nyquist point. For the predicted parameter values of Equation 22, limit cycles may occur since the open-loop transfer function $G(s).H(s)$ intersects the $-1/N_A$ function (Figures 6b-6d). For the parameter value in Figure 6e, the system is possibly unstable due to an encirclement of the Nyquist point.

Simulations were generated in Simulink\textsuperscript{TM} to crosscheck with the analytical results. The solver was set to use the Dormand-Prince method, a common integration algorithm for solving differential equations, and to run for 4000 iterations. Before introducing the exogenous input, we tested the initial and final steady state conditions of the variables both analytically through the Initial and Final Value Theorems and via simulation by running it until steady state is reached. Figure 7 illustrates the order rate responses, $do(t)$ and $o(t)$, corresponding to the Nyquist diagrams of Figure 6 and given an unit step input demand $d = 1$. In stable systems (Figure 7a), when the reference input signal is a constant (step input), the output signal is a constant in steady-state. However, the figure illustrates that limit cycles do occur for the estimated parameter values (Figure 7b-7d). We also compare the estimated and simulated values of $\omega_{do}$ and $A_{do}$. As the non-negativity constraint becomes more effective and $N_{A_{do}}$ decreases, predicted values become less accurate. This can be explained by the asymmetric nature of this nonlinearity.
Figure 6: Nyquist plots for the ordering rule in the nonlinear APIOBPCS
Figure 7: Order rate responses to a constant demand input
After confirming the existence of limit cycles, we should now determine their stability. The question of limit cycle stability is posed in terms of its behaviour following amplitude and/or frequency disturbances. If the limit cycle returns to its original equilibrium state it will be called stable, whereas if either its amplitude or frequency changes to another equilibrium state, it will be called unstable. We use Loeb’s criterion (1956) to determine the stability of the limit cycle found in the ordering system (see Appendix A).

Previously, we determined $\omega_{do}$ and $A_{do}$ of the limit cycle and that the following identity holds:

$$1 + N_{A_{do}}G(j\omega_{do})H(j\omega_{do}) = 0. \quad (27)$$

The explicit complex form of this equation can be obtained by expressing the quantities $N_{A_{do}}$, $G(j\omega)$ and $H(j\omega)$ of Equations 8, 24 and 25, respectively, in terms of their real and imaginary parts and making $s = j\omega$, resulting in:

$$\pi T_i T_p T_w \omega^2 + T_w \gamma - T_w \cos(\gamma) \sin(\gamma) + j(T_i T_p \omega \gamma - T_i T_p \omega \sin(\gamma) - \pi T_i T_w \omega) = 0. \quad (28)$$

According to Loeb’s criterion (detailedly presented in Appendix A), for the limit cycle equilibrium condition to be stable it is necessary that

$$\frac{\partial U}{\partial A} \frac{\partial V}{\partial \omega} - \frac{\partial U}{\partial \omega} \frac{\partial V}{\partial A} > 0, \quad (29)$$

where $U$ and $V$ are the real and imaginary parts of Equation 28, respectively. Replacing the value of $\gamma$ in Equation 28 and taking the required four partial derivatives at $(A_{do}, \omega_{do})$, results in

$$\frac{\partial U}{\partial A} \bigg|_{A_{do}, \omega_{do}} = \frac{2B_{do} \sqrt{1 - \frac{B_{do}^2}{A_{do}^2} T_w}}{A_{do}^2}, \quad \frac{\partial V}{\partial \omega} \bigg|_{A_{do}, \omega_{do}} = \frac{2B_{do} \sqrt{1 - \frac{B_{do}^2}{A_{do}^2} T_i^3}}{A_{do}^2 T_w},$$

$$\frac{\partial U}{\partial \omega} \bigg|_{A_{do}, \omega_{do}} = -2\pi \sqrt{-T_i T_w^3}, \quad \frac{\partial V}{\partial A} \bigg|_{A_{do}, \omega_{do}} = \frac{B_{do} \sqrt{1 - \frac{B_{do}^2}{A_{do}^2} T_i T_p}}{A_{do}} - \pi T_i T_w + T_i T_p \cos^{-1} \left( -\frac{B_{do}}{A_{do}} \right). \quad (30)$$
Inserting these values into Equation A.7 reveals, as the requirement for a stable oscillation, the conditions that $A_{do} > B_{do}$ and that $-T_p \leq T_w \leq 0$ must be satisfied. Since no limit cycle occurs for $A_{do} < B_{do}$ (the system behaves linearly, see Figure 2a) and $-0.5T_p \leq T_w \leq 0$ (see Equation 22), all limit cycles resulting from the nonlinear production and inventory control system are stable. We can crosscheck these results graphically by examining Figure 6 again. In the limit cycle cases (Figures 6b-6d), an increase in the amplitude $A_{do}$ would correspond to a shift to the left of the $-\frac{1}{N_{1_{do}}}$ point in the s-plane. With that change, $G(s)H(s)$ would not encircle the critical point and therefore $A_{do}$ would revert back to its original state. Conversely, a negative perturbation in $A_{do}$ leads to an unstable system configuration which causes $A_{do}$ to increase and converge back to its original equilibrium point.

5.4. Linear compensation to control limit cycles

In this section, we explore linear loop compensation techniques to re-design the nonlinear system under threat of limit cycling behaviour so that steady-state and transient response requirements are met. We chose the two most established methods: lead/lag and pure compensations. The method of choice is frequently critically dependent on the exact details of the linear and nonlinear elements included in the loop (Gelb and Velde, 1968). In hard systems, the system designer is frequently constrained to place a compensation network at locations where the feedforward (an anticipatory response to expected changes in demand) or feedback paths are physically accessible. In our case, this is simply done by inserting an equation in the ordering rule and we chose to introduce it in the feedforward path before the nonlinear element, as demonstrated in Figure 8.

Figure 8: Simplified and compensated block diagram for the ordering rule in the nonlinear APIOBPCS

\[
\begin{align*}
&d \\
&\rightarrow T_w + s(T_i T_p + T_a T_w + T_i T_w) \\
&\rightarrow \frac{s^2 T_a T_i T_w + s T_i T_w}{s^2 T_a T_i T_w + s T_i T_w} \\
&\rightarrow L(s) \\
&\rightarrow (1 + s T_a) \left( s T_i T_p + s T_w \right) \\
&\rightarrow \frac{(1 + s T_u) (s T_i T_p + T_w)}{(1 + s T_p) (s T_i T_p + s T_a T_w + s T_i T_w + T_w)}
\end{align*}
\]
A lead or lag compensation element of transfer function

\[ L_1(s) = \frac{\beta \tau s + 1}{\tau s + 1} \]

lead network for \( \beta > 1 \), lag network for \( \beta < 1 \)  \hspace{1cm} (31)

can be introduced to attenuate signals with high frequencies (relative to \( 1/\beta \tau \)).

Figure 9 illustrates the lead and lag compensation strategies for a fixed time constant, \( \tau = 1 \), in comparison with the uncompensated system, whose parameter settings correspond to the ones entered in Figure 7b. The figure depicts the uncompensated loci \( G(j\omega)H(j\omega) \) intersecting the \( -1/N_{A_{do}} \) when the phase is -180 degrees and amplitude is \( A_{do} \approx B_{do} \). The frequency of oscillation, \( \omega_{do} \) is marked by a red star in the uncompensated and compensated models. The amplitude-phase plane indicates that limit cycles will continue to exist for lag compensations, but their frequencies will be lower and amplitudes higher. Lead compensations can move the system out of the limit cycling zone, for instance, when \( \beta = 1.2 - 2.5 \). However, when \( \beta \) values are too large, limit cycles of higher amplitudes and lower frequencies will occur (see the insert of Figure 9 for a zoomed in view). In Figure 10, we present the corresponding order rate responses. For the uncompensated order rate response, the reader can refer back to Figure 7b. For the lag compensation (Figures 10a and 10b), intrinsic oscillations persist, but as \( \beta \) increases the amplitude of oscillation decreases and its frequency

![Figure 9: Amplitude-phase plane: effects of linear lead and lag compensation](image_url)
Figure 10: Order rate responses after lead and lag compensation
increases. Lead compensation can be used to eliminate limit cycles (Figures 10c, 10d and 10e) but its setting should be carefully chosen to not return responses to exhibit oscillatory behaviour (Figure 10f).

Another compensation method that can be used is the pure gain compensation in which a linear pure gain $k$ is inserted in the feedforward path just before the nonlinear element, such as

$$L_2(s) = k$$ higher-than-unit gain for $k > 1$, lower-than-unit gain $k < 1$. \hfill (32)

The effect of $k$ is simply to raise or lower a frequency response locus on the amplitude-phase plane. This operation will not affect the response frequency since the nonlinearity in the ordering system is approximated by a non-phase-shifting describing function. Figure 11 illustrates the amplitude-phase plane for pure gain compensation when values of $k$ are less than and greater than 1. Limit cycles can be easily eliminated by inserting a lower-than-unity gain. When using a $k$ value greater than 1, the system will continue to oscillate at the same frequency $\omega_{do}$ but at higher amplitude. Figure 12 illustrates the effectiveness of this compensation method on controlling oscillations in the order rate when $k < 1$ (Figures 12c and 12d). As predicted, limit cycles continue to occur for $k > 1$ with same frequency and increased amplitude (Figures 12a and 12b).

![Figure 11: Amplitude-phase plane: effects of pure gain compensation](image-url)
6. Discussion and conclusion

This paper proposes a methodological framework that can be used to investigate the occurrence of sustained oscillations in production and inventory control models due to the presence of nonlinearities. Using a benchmark model, we demonstrated how the describing function technique can be used to analytically identify the existence of limit cycles, their characteristics, such as amplitude, frequency and stability, and consequently to devise a suitable compensation method to control such oscillations. In our model, two nonlinearities were investigated but when it comes to limit cycles, only the non-negativity constraint is of concern for the supply chain designer. Uninformed replenishment policies and some aperiodic configurations combined with changes in the lead-time can cause the system to have an
oscillatory behaviour, which can in turn drive costs up and service levels down. In practice, an example of uninformed replenishment policy was evidenced by Spiegler et al. (2016b), who noticed that the replenishment system of a DC in a grocery retailer was set to have WIP target values equal to zero, causing inventory offsets. Although in this example, this uninformed policy did not cause periodicity in the supply chain system (rogue seasonality), it contributed to the bullwhip and ripple effect.

The proposed method can be used for investigating the root causes of oscillations in other nonlinear production and inventory control systems. It enables an understanding of how the amplitude and frequency of oscillations are affected by different signals passing through the nonlinearity. After insights are gained about the effect of the nonlinearity, suitable compensation methods can be used to avoid and manage limit cycles. In comparison to a simulation-only approach and using linear control theory, our method enables greater insights to be gained regarding more realistic production planning and inventory management model.

Several interesting observations and practical implications can be inferred from this study:

1. The describing function method draws attention to the importance of being aware of the frequency and amplitude of demand signals. The non-negativity constraint in the ordering rule makes the order rate reach zero more often when demand variability is high, which may skew perceptions of demand even further. Although this is a rather intuitive finding, we are able to find the percentage of non-zero orders via the describing function gain given the demand amplitude and mean. Hence, depending on the demand characteristics, appropriate levelling strategies may be implemented (Section 4.1).

2. Since the nonlinearity in the shipment system is frequency-dependent, we find that not only will high demand levels (in fact amplitude) trigger this constraint but so too will medium-low frequency demands, possibly causing more backlogs. Hence, supply chain managers should create incentives to increase demand frequencies. This strategy is also supported by Wang et al. (2015b), who found that low ordering frequencies can increase the bullwhip, and Spiegler et al. (2016b) when looking at the supply chain resilience of a grocery retailer. Managing demand by encouraging more frequent sales is industry dependent. For instance, grocery customers’ normal shopping habits tend to

---

1 Standard deviation is function of amplitude in sinusoidal waves
generate high-frequency demands, with peaks concentrated on Fridays and weekends when customers tend to buy most of the groceries needed for the week (Spiegler et al., 2016b). However, clothing and toy industries have longer seasonality periods (Wong et al., 2006), therefore these companies would benefit from demand shaping strategies, such as new product launch, price management, marketing and advertising. (Section 4.2).

3. The nonlinearity in the shipment system has no impact on the order rate and WIP (refer to Equations 17 and 19). This is due to the fact that this is a physical, rather than an information, constraint. Moreover, this nonlinearity does not cause oscillatory behaviour in the system. Hence, the supply chain designer does not have to be concerned with this capacity limitation when setting parameter values (Section 5.2).

4. The non-negativity nonlinearity can cause limit cycles, which are oscillations intrinsic to the nonlinear production and inventory control system itself and not imposed by the demand. This finding is in line with the work of Wang and Disney (2012) and Wang et al. (2014). However, our framework was able to predict not only when these oscillations may occur but also their amplitude, frequency and stability. All limit cycles resulting from the nonlinearity in the ordering rule are stable, which means that they return to their original equilibrium state after being perturbed and responses will not become unstable when the system is no longer autonomous (demand is random). As the nonlinearity becomes more effective, the amplitude and frequency of order rate oscillations increase and decrease, respectively. This problem may be exacerbated as this information signal propagates upstream in the supply chain. This information is vital for the supply chain designer to distinguish between exogenous- and endogenous-related oscillations and to devise compensation methods when necessary (Section 5.3).

5. We explore the two most common compensation methods to control oscillations. These compensation techniques can be inserted into the system to manage and avoid oscillations induced by nonlinearities. We found that lead and lower-than-unit pure gain compensations are better suited for overcoming limit cycling behaviour in our benchmark model. For the pure gain compensation method, the supply chain designer benefits from its damping effect on order rate, therefore decreasing both oscillations and demand amplification. On the other hand, lead compensations enable adjustment of
the response time constant, which in turns determines the speed with which the system can respond to changes in demand (Section 5.4).

Our research serves as a cornerstone for designing optimal control policies and rules in nonlinear supply chain systems, providing a continuation of recent research on the use of modern control theory (Wang and Disney, 2012; Wang et al., 2014, 2015b; Spiegler et al., 2016a,b). The research can be extended further by assessing the impact of other nonlinearities in the ordering rule, such as fixed and variable production capacity, and evaluating how robust the compensation methods are under changing constraints. This work is limited to the dynamics of a single-echelon supply chain system, but the proposed framework could also be applied to analyse nonlinear dynamics of multi-echelons and other nonlinear supply chain structures beyond informational, such as functional, organisational and financial, to accomplish multi-structural supply chain management.

An unconventional direction for future research would be the study of deliberately setting production and inventory control systems to assume limit cycling behaviour and explore filter theory to level orders of different customers in a vertical supply chain management network structure, for example. Since the amplitude and frequency of oscillations in the order rate can be controlled, the system can be set to produce anti-phase signals to filter responses from the manufacturer. Towill and Vecchio (1994) had already explored filter theory by varying the different echelon dynamics in order to minimise stockholding while maintaining service levels. Given the risk associated with this research idea, an in-depth robustness analysis should be undertaken especially under the effect of lead-time disturbances. Moreover, a combination of compensation networks can be explored to produce robust control designs that fit supply chain characteristics and strategies.

Acknowledgements

The authors would like to thank Brunel Business School for providing a writing retreat that helped bring this work to fruition.

References


Appendix A. Loeb’s criterion for stability determination of limit cycles

Here we present the Loeb’s criterion (Loeb, 1956) in more detail by following the discussion of Gelb and Velde (1968). Previously we have seen that when determining $\omega_{do}$ and $A_{do}$ of an equilibrium limit cycle state, the following identity holds:

$$1 + N_{A_{do}}G(j\omega_{do}).H(j\omega_{do}) = 0. \quad (A.1)$$

The complex form of A1 can be obtained by expressing the quantities $N_{A_{do}}$, $G(j\omega_{do})$ and $H(j\omega_{do})$ in terms of their real and imaginary parts, resulting in:

$$U(A_{do}, \omega_{do}) + jV(A_{do}, \omega_{do}), \quad (A.2)$$

where the definitions of $U$ and $V$ are the real and imaginary parts, respectively. If we allow small perturbations in the limit cycle amplitude, rate of change in amplitude and frequency by introducing the following changes in A2:

$$A_{do} \rightarrow A_{do} + \Delta A$$
$$w_{do} \rightarrow w_{do} + \Delta w + j\Delta \sigma \quad (A.3)$$

The perturbation in the rate of change of amplitude has been associated with the frequency term. That is $\Delta \sigma = \dot{A}/A$. Hence, we have

$$\frac{\partial U}{\partial A} \Delta A + \frac{\partial U}{\partial \omega} \Delta \omega + \frac{\partial V}{\partial \omega} (\Delta \omega + j\Delta \sigma) + j \frac{\partial V}{\partial A} \Delta A = 0 \quad (A.4)$$

Satisfying this equation requires that its real and imaginary parts separately disappear.

$$\frac{\partial U}{\partial A} \Delta A - \frac{\partial U}{\partial \omega} \Delta \sigma + \frac{\partial U}{\partial \omega} \Delta \omega = 0$$
$$\frac{\partial V}{\partial A} \Delta A + \frac{\partial U}{\partial \omega} \Delta \sigma + \frac{\partial V}{\partial \omega} \Delta \omega = 0 \quad (A.5)$$
Eliminating $\Delta \omega$ from this set of equations yields the following single relationship

$$\left( \frac{\partial U}{\partial A} \frac{\partial V}{\partial \omega} - \frac{\partial U}{\partial \omega} \frac{\partial V}{\partial A} \right) \Delta A = \left[ \left( \frac{\partial U}{\partial w} \right)^2 + \left( \frac{\partial V}{\partial w} \right)^2 \right] \Delta \sigma. \quad (A.6)$$

For a limit cycle to be stable, a positive increment $\Delta A$ must cause a negative $\Delta \sigma$. Hence, for the proposed limit cycle equilibrium condition to be stable, it is necessary that

$$\frac{\partial U}{\partial A} \frac{\partial V}{\partial \omega} - \frac{\partial U}{\partial \omega} \frac{\partial V}{\partial A} > 0. \quad (A.7)$$