Sustained oscillations via coherence resonance in SIR

Rachel Kuske*, Luis F. Gordillo †, Priscilla Greenwood†

September 26, 2006

Abstract

Sustained oscillations in a stochastic SIR model are studied using a new multiple scale analysis. It captures the interaction of the deterministic and stochastic elements together with the separation of time scales inherent in the appearance of these dynamics. The nearly regular fluctuations in the infected and susceptible populations are described via an explicit construction of a stochastic amplitude equation. The agreement between the power spectral densities of the full model and the approximation verifies that coherence resonance is driving the behavior. The validity criteria for this asymptotic approximation give explicit expressions for the parameter ranges in which one expects to observe this phenomenon.

1 Introduction

Sustained nearly-regular oscillations in large populations infected with such diseases as measles, chickenpox and flu have been a subject of puzzlement for many years. Often such diseases are modeled by a deterministic mean-field approximation described by differential equations which have a stable positive equilibrium state if the basic reproductive number $R_0$ is larger than 1. It follows that in the phase space of the system the trajectories spiral in to the endemic point. In contrast, it has been observed in Monte Carlo simulations ([1], [2], [3] and related stochastic realizations as shown in Figure 1), that the path of a stochastic susceptible-infective-recovered model (SIR) or similar stochastic model with a large population, may follow the damped trajectory of the corresponding deterministic model for a certain initial time, after which the stochastic path remains oscillatory. These oscillations have a narrow frequency distribution, evidenced by the power spectral density (PSD) of the process [4], [5], [6], and a stochastically varying amplitude. This phenomenon, in which stochastic fluctuations sustain nearly-periodic oscillations in a system which has a stable constant equilibrium in the deterministic limit, has been called coherence resonance or autonomous stochastic resonance [4, 6, 7].

Oscillations which do not damp, or damp very weakly, can be produced in deterministic epidemic models in several ways, essentially by making the model sufficiently complex. Sustained oscillations can be produced by adding seasonality,
quarantine, age structure, or multiple strains of infection in various combinations to the model ([8], [9], [10] [11]). Our analysis identifies the mechanism by which nearly periodic oscillations are sustained in the stochastic model, even without these additional factors. Both deterministic and stochastic characteristics persist via the multiple time scales inherent in the model, as shown by an approximation based on stochastically modulated regular oscillations. Parameter ranges for observing this behavior can be found directly from our explicit construction. In this paper we consider the simple SIR model, in order to illustrate the approach and to highlight the fundamental characteristics leading to sustained oscillations without the distractions of other complicating factors.

In a deterministic SIR model with demography there is a susceptible population of size $S$, an infective population of size $I$ and a population of recovered individuals of size $R$. The populations evolve in time according to the rates

\begin{align*}
\text{transition} & \quad \text{rate} \\
S & \rightarrow S + 1 \quad \mu N \\
S & \rightarrow S - 1 \quad \beta SI / N + \mu S \\
I & \rightarrow I + 1 \quad \beta SI / N \\
I & \rightarrow I - 1 \quad (\gamma + \mu)I \\
R & \rightarrow R + 1 \quad \gamma I \\
R & \rightarrow R - 1 \quad \mu R
\end{align*}

(1)

where $N$ is fixed and equal to the total population size. The rate of birth or death per individual is $\mu$, $\gamma$ is the recovery rate and $\beta I / N$ is the average number of contacts with infectives for one susceptible per unit time. The basic reproductive number, $R_0$, is the number of new infectives produced by one infective introduced in a completely susceptible population, (see [12]), and is defined by

$$R_0 = \frac{\beta}{\mu + \gamma}.$$  

(2)
The deterministic model can be written as

\[
\frac{dS}{dt} = \mu(N - S) - \beta \frac{SI}{N},
\]

(3)

\[
\frac{dI}{dt} = \beta \frac{SI}{N} - (\gamma + \mu)I,
\]

the equation for \( R \) being redundant, with \( S + I + R = N \). If \( R_0 > 1 \) there is a stable endemic \((I = 0)\) equilibrium.

The corresponding stochastic model is a continuous time Markov process \((S_t, I_t, R_t) : t \in [0, \infty)\) with state space \(Z_3^+\). The parameter \( N \) is the expected population size. The rates in (1) become the conditional transition rates of the stochastic process \((S; I; R)\), that is,

\[
P(S_t + t = s + 1 | S_t = s) = N t + o(t), \text{ etc.}
\]

We start the process with \( S + I + R = N \) and the expected sum remains \( N \). The states with no infectives are absorbing, the other states being transient.

We express the stochastic equations of the process in a form easily compared with the equations (3) of the deterministic model. To each increment we add and subtract its conditional expectation, conditioned on the value of the process at the beginning of the time increment of length \( \Delta t \). Each increment of the process is then decomposed into the sum of the expected value of the increment and a sum of centered increments [13].

\[
\Delta S = (\mu(N - S) - \beta \frac{SI}{N}) \Delta t + \Delta Z_1 - \Delta Z_2,
\]

(4)

\[
\Delta I = (\beta \frac{SI}{N} - (\gamma + \mu)I) \Delta t + \Delta Z_2 - \Delta Z_3.
\]

The mean of \( \Delta S = S_{t+\Delta t} - S_t \) is \((\mu N - \mu S_t - \beta S_t I_t/N)\Delta t\). We write the centered increment \( \Delta S - E[\Delta S] \) as the difference of two increments \( \Delta Z_1 - \Delta Z_2 \). The increment \( \Delta Z_1 \) is the difference of the two centered Poisson increments corresponding to births into the class \( S \) and deaths out of \( S \). It has zero mean and the variance is the sum of two variances, \( \mu(N + S)\Delta t \). We write the centered Poisson increment \( \Delta Z_2 \), corresponding to infections, separately because this same increment appears in the corresponding representation of \( \Delta I \) with opposite sign. The increments \( \Delta Z_1, \Delta Z_2, \Delta Z_3 \) are independent with variances \( \mu(N + S)\Delta t, \beta \frac{SI}{N}\Delta t \) and \((\gamma + \mu)I\Delta t\), respectively.

In developing the theory we replace the conditionally centered Poisson variables \( \Delta Z_i \) in (4) by increments of Brownian motion, \( dW_i \), with the same standard deviations, and introduce rescaled, dimensionless variables. For small \( \Delta t \) and large \( N \), the process with centered Poisson increments is well-approximated by the process with Gaussian increments with the same conditional variances. Note that (4) does not converge to a diffusion as \( N \to \infty \), but rather when normalized it converges to a deterministic system. In this large \( N \) limit, the random fluctuations \( \Delta Z_j \) appear as higher order corrections to the deterministic dynamics; nevertheless, they can still play a significant role. This modeling approach is similar in some respects to the development of stochastic differential equation models in random molecular reactions [14] and approximations of shot noise [15].

Aparicio and Solari [3] consider a model similar to ours and give an explanation of the non-damping of the stochastic (Poisson) version. They define a parabola in phase space delineating positive (negative) average change in a Liapunov function, \( E \), for the process inside (outside) of the parabola. This heuristic argument suggests
that the oscillations of the stochastic system do not die out regardless of the initial state, as is observed more generally for systems with contracting drift and either additive or multiplicative noise [16]. McKane and Newman [5] have observed oscillations in stochastic predator-prey models, which are structurally similar to SIR. They used the van Kampen expansion based on the linearized system to obtain a Gaussian approximation for the probability density, numerically computing the PSD which showed a strong peak near the frequency of the deterministic system. However, neither of these studies explains the nearly regular behavior of the oscillations, demonstrated by the peak in the PSD and directly related to the multiscale structure of the dynamics. Näsell [17] computes an approximation to the asymptotic distribution at one time point of the stochastic SIR, conditioned to stay positive. But this distribution contains no information about the character of the oscillations or the asymptotic dynamics of the process at large times, which we study here.

Our construction of the solution captures the oscillatory behavior explicitly, and compares well with the PSD of the full system for a significant range of parameters. We derive stochastic amplitude equations for the envelope of the nearly regular oscillations, enabling the identification of both the stochastic and deterministic elements of the dynamics. Stochastic multi-scale analysis exploits the separation of time scales, which we show is an inherent system property over a significant parameter range and is directly related to the narrow distribution of frequencies in the PSD, typical of coherence resonance. The validity criteria for this asymptotic approximation give explicit expressions for the parameter ranges in which this effect exists.

Coherence resonance has been demonstrated in a number of other models, where noise sustains oscillations near a certain frequency in otherwise quiescent systems. This phenomenon was reported by Gang et al. in [4], who studied a two dimensional model in which a limit cycle can be generated or eliminated by adjusting a control parameter. When this parameter is just below the critical value, where no coherent oscillation persists without noise, the introduction of noise stimulates a coherent motion whose frequency peak depends on the noise variance. It has also been observed in a large number of studies of noise-induced synchrony in networks of coupled excitable systems, including coupled integrate-and-fire models [19], FitzHugh-Nagumo models [7, 20], Hodgkin-Huxley models [21], and bursting models [22]. Optimal coherence at a finite noise level was explained by Wiesenfeld [23] who revealed how the noise controls the structure of the power spectrum. The same phenomenon is studied analytically by Yu et al. [6] for a canonical model near a Hopf bifurcation with added noise, where a multiscale method is used to derive the stochastic amplitude and phase dynamics. The underlying model in the SIR case does not have a Hopf bifurcation; in fact, in the biologically relevant parameter regime it has no bifurcation. Nevertheless, the same multi-scale mechanism is responsible for sustaining the oscillations, as revealed by the analysis below.

A number of stochastic models have the additional property that fluctuations eventually drive the process to extinction. Various approaches have been used to focus on the long term behavior while avoiding extinction. For example, Näsell [17] considers the quasi-stationary distribution of the process conditioned not to hit the absorbing states. A second strategy, adopted for example in [3] and [1], is to replace the $\Delta Z_i, i = 1, 2, 3$ by Gaussian increments and ignore the fact that in the first
few orbits of the endemic point, the process passes close to the absorbing set. In the following sections we study the phenomenon in regions of the parameter space where the absorbing state is relatively insignificant; either the system is unlikely to approach extinction or recurrent outbreaks could be produced by random reintroduction of infectives, as is appropriate for larger populations. This viewpoint on the stochastic model goes back to Bartlett [24], who also observed that stochastic fluctuations play an important role in the qualitative nature of recurrent epidemic dynamics for finite population sizes.

In Section 2 we introduce the dimensionless linearized system and identify the slow time scale, directly responsible for the coherence resonance. In Section 3, we apply the stochastic multi-scale method, which separates the stochastic amplitude of the oscillations from the deterministic frequency corresponding to the PSD peak. We identify the stochastic modulating process which gives the amplitude of the oscillations. Its variance is proportional to the variance of the noise and to the inverse of the slow time scale. This construction demonstrates that the fluctuations are amplified according to the modulating process. In Section 4 we give the range of parameters for which this phenomenon is observed.

2 The dimensionless linearized problem

We consider the system

\[
\begin{align*}
    ds &= \left( \mu(N - S) - \frac{\beta}{N}SI \right) dt + G_1 dW_1 - G_2 dW_2, \\
    di &= \left( \frac{\beta}{N}SI - (\gamma + \mu)I \right) dt + G_2 dW_2 - G_3 dW_3, \\
    G_1 &= \sqrt{\mu(N + S)}, \quad G_2 = \sqrt{\frac{\beta}{N}SI}, \quad G_3 = \sqrt{(\gamma + \mu)I}.
\end{align*}
\]

If \( R_0 > 1 \), the deterministic system obtained by omitting the \( dW_i \) has a unique nontrivial stable equilibrium point \((S_{eq}, I_{eq})\) at

\[
S_{eq} = \frac{N}{R_0}, \quad I_{eq} = \frac{N\mu}{\beta}(R_0 - 1).
\]

We introduce the dimensionless variables

\[
u = \frac{S - S_{eq}}{S_{eq}}, \quad \frac{I - I_{eq}}{I_{eq}},
\]

and rescale time from \( t \) to \( \Omega t \), where \( \Omega = \sqrt{\frac{\beta\mu}{R_0}(R_0 - 1)} \). This choice of \( \Omega \) yields damped oscillations with unit frequency for the rescaled system in the absence of
noise. The stochastic system (5) becomes

\[
du = \frac{1}{\Omega} \left[ (-\mu - \beta I_{eq}) u - \frac{\beta I_{eq}}{N} v - \frac{\beta I_{eq}}{N} uv \right] dt + \sqrt{\frac{\mu}{\Omega S^2_{eq}}} (N + S_{eq}(u + 1)) dW_1(t) - \sqrt{\frac{\beta I_{eq}}{\Omega N S_{eq}}} (v + 1)(u + 1) dW_2(t),
\]

\[
dv = \frac{1}{\Omega} \left[ \frac{\beta S_{eq}}{N} (u + v) + \frac{\beta S_{eq}}{N} uv - (\gamma + \mu)v \right] dt + \sqrt{\frac{\beta S_{eq}}{\Omega N I_{eq}}} (v + 1)(u + 1) dW_2(t) - \sqrt{\frac{\gamma + \mu}{\Omega I_{eq}}} (v + 1) dW_3(t).
\]

(7)

We show below that the noise in (7) is small for a significant range of parameters. Consequently one would expect the dynamics to be concentrated around \( u = 0 \) and \( v = 0 \), the stable equilibrium for the deterministic version of the system. Then it is appropriate to linearize (7) around \( u = 0 \) and \( v = 0 \) to obtain an approximation for \( u_1 \) and \( v_1 \). We also set \( u = v = 0 \) in the diffusion coefficients, as a leading order approximation to the diffusion, which yields

\[
d \begin{pmatrix} u \\ v \end{pmatrix} = M \begin{pmatrix} u \\ v \end{pmatrix} dt + G \begin{pmatrix} dW_1 \\ dW_2 \\ dW_3 \end{pmatrix},
\]

(8)

\[
M = \begin{pmatrix} -\frac{\mu R_0}{\beta I_{eq}} & -\frac{\mu R_0 - 1}{\beta I_{eq}} \\ \frac{\beta S_{eq}}{\Omega N S_{eq}} & 0 \end{pmatrix},
\]

(9)

\[
G = \begin{pmatrix} \sqrt{\frac{\mu}{\Omega S^2_{eq}}} (N + S_{eq}) & -\sqrt{\frac{\beta I_{eq}}{\Omega N S_{eq}}} & 0 \\ 0 & \sqrt{\frac{\beta S_{eq}}{\Omega N I_{eq}}} & -\sqrt{\frac{\gamma + \mu}{\Omega I_{eq}}} \end{pmatrix},
\]

(10)

\[
= \begin{pmatrix} g_1 & -b^2 g_2 & 0 \\ 0 & g_2 & -g_2 \end{pmatrix}.
\]

We have used the definitions of \( R_0 \) and \( I_{eq} \) given by (2) and (6) and defined \( b^2 = I_{eq}/S_{eq} \). The noise in (8) is interpreted in the Ito sense.

Solutions of the deterministic version of (8) are given in terms of the eigenvalues of \( M \), which are

\[
\lambda = -\epsilon^2 \pm \sqrt{\epsilon^4 - 1},
\]

where

\[
\epsilon^2 = \frac{\mu R_0}{2\Omega}.
\]

For \( \epsilon \ll 1 \) the solution of the deterministic problem near the equilibrium is approximated by

\[
\begin{pmatrix} u \\ v \end{pmatrix} \sim C_1 e^{-\epsilon^2 t} \begin{pmatrix} b \cos t \\ b \sin t \end{pmatrix} + C_2 e^{-\epsilon^2 t} \begin{pmatrix} b \sin t \\ -b \cos t \end{pmatrix},
\]

(11)

if we drop \( O(\epsilon^2) \) corrections to the frequency. That is, for \( \epsilon \ll 1 \) oscillations with unit frequency with slowly decaying amplitude are observed. In Figure 2 we show that \( \epsilon^2 \) is small over a significant range of parameters. Furthermore, in these parameter regimes the diffusion coefficients \( g_j \) for \( j = 1, 2 \) are often \( O(\epsilon) \) or smaller,
as discussed in Section 4. Observing that (8) is essentially the linearized version of the deterministic problem, one might expect that qualitatively similar behavior could be obtained by adding arbitrary white noise to the deterministic problem, and then linearizing. This is indeed true; however, in such a model the diffusion coefficients would not be necessarily related to the system parameters. In the next section we apply a multi-scale approximation for the linearized system with additive noise (10) in these parameter regions, which yields the leading order behavior of the oscillations of the stochastic system (5). In Section 5 we discuss methods for including higher order corrections in the nonlinear problem with both additive and multiplicative noise.

3 Multiscale Analysis of the Linearized Stochastic Model

The behavior of the solution to the linearized deterministic problem (11) suggests a form of solution to (8) with oscillations on the $t$ scale and with modulations described by a slowly varying amplitude or envelope (cf. [25]). We might expect that the slow variation of the envelope would be destroyed by the stochastic fluctuations, but we show through the method of multiple scales that, in fact, for relatively small noise, this slow variation survives. The resonance of the stochastic variations with the frequency of the deterministic system leads to an amplification of the otherwise damped oscillations with a slowly varying stochastic modulation. Below we give an approximation in a form which explicitly captures both the multiple scales behavior and the resonance with the deterministic frequency. While it is true that one could use a standard approach to solve the linearized stochastic system 8 and obtain the PSD of the resulting process [15], such an approach will not reveal the underlying multi-scale structure. Instead the slow variation and the resonance effect are immediately apparent from our multi-scale analysis. Furthermore, the approach presented here has been used more generally in stochastic nonlinear models [18, 28].

We consider a proposed approximation for the solution of the linearized stochastic model (8),

$$\begin{pmatrix} u \\ v \end{pmatrix} = A(T) \begin{pmatrix} b \cos t \\ \sin t \end{pmatrix} + B(T) \begin{pmatrix} b \sin t \\ -\cos t \end{pmatrix},$$  

(12)

where $A(T)$ and $B(T)$ incorporate both the slow decay and the stochastic part of the solution, and where $T$ varies slowly relative to $t$, $T = \epsilon^2 t$, $\epsilon \ll 1$. The form of (12) is commonly used in multiple scale analyses to derive amplitude or evolution equations for deterministic systems in parameter regimes near a bifurcation point [25, 26]. The difference here is that $A$ and $B$ must capture the stochastic behavior. Since the stochastic nature of the model does not allow a standard application of the multiple scales expansion, we use a modified approach which incorporates Ito calculus and the properties of the noise [27].

Following the substitution of (12) into (8), the stochastic system becomes

$$\begin{align*}
\frac{du}{dt} &= (-2b^2[ A \cos t + B \sin t ] + b[-A \sin t + B \cos t ] )dt + g_1dW_1 - b^2g_2dW_2, \\
\frac{dv}{dt} &= [ A \cos t + B \sin t ]dt + g_2dW_2 - g_2dW_3 
\end{align*}$$  

(13)
Since the amplitude functions $A$ and $B$ are stochastic coefficients in the proposed form (12) which should approximately solve the diffusion system (8), we suppose they are driven by Brownian motions. Given the relatively simple form of (12), a reasonable form for the amplitude equation is that of a diffusion with constant diffusion coefficients,

$$
\begin{pmatrix}
  dA \\
  dB
\end{pmatrix} = \begin{pmatrix}
  f_1(A, B) \\
  f_2(A, B)
\end{pmatrix} dT + \begin{pmatrix}
  dN_1(T) \\
  dN_2(T)
\end{pmatrix}.
$$

(14)

For convenience, we may write the noise terms, $N_1$ and $N_2$, as

$$
\begin{pmatrix}
  dN_1(T) \\
  dN_2(T)
\end{pmatrix} = \Sigma_1 \begin{pmatrix}
  d\xi_{11}(T) \\
  d\xi_{12}(T)
\end{pmatrix} + \Sigma_2 \begin{pmatrix}
  d\xi_{21}(T) \\
  d\xi_{22}(T)
\end{pmatrix} + \Sigma_3 \begin{pmatrix}
  d\xi_{31}(T) \\
  d\xi_{32}(T)
\end{pmatrix},
$$

(15)

where the $\xi_{ij}$’s are independent standard Brownian motions on the slow time-scale $T$ and the $\Sigma_j$’s are constant matrices. This is a general form motivated by the form of the noise in (8), and we shall see some simplifications later on.

The drift coefficients, $f_1, f_2$ and the diffusion coefficients $\Sigma_1, \Sigma_2, \Sigma_3$ appearing in (14) are unknowns and must be derived by relating (14) to (13). For a solution $(u, v)'$ of the form (12), Ito’s formula can be used to obtain a second expression for $du$ and $dv$, of the form,

$$
\begin{align*}
du &= \frac{\partial u}{\partial A} dA + \frac{\partial u}{\partial B} dB + \frac{1}{2} Q_1 \frac{\partial^2 u}{\partial A^2} dt + Q_2 \frac{\partial^2 u}{\partial A \partial B} dt + \frac{1}{2} Q_3 \frac{\partial^2 u}{\partial B^2} dt + \frac{\partial u}{\partial t} dt, \\
& \quad \text{(16)}
\end{align*}
$$

and similarly for $dv$ with $u$ replaced by $v$ in (16), with $Q_k$ in terms of the entries of the diffusion coefficients $\Sigma_j$ in (15). Since the proposed form (12) is linear in $A$ and $B$, the second derivatives vanish in (16). Using (14), the $u$ and $v$ equations are then

$$
\begin{pmatrix}
  du \\
  dv
\end{pmatrix} = \begin{pmatrix}
  b \cos t & b \sin t \\
  \sin t & -\cos t
\end{pmatrix} \begin{pmatrix}
  f_1 \\
  f_2
\end{pmatrix} dT + \begin{pmatrix}
  -bA \sin t + Bb \cos t \\
  A \cos t + B \sin t
\end{pmatrix} dt \\
+ \begin{pmatrix}
  b \cos t & b \sin t \\
  \sin t & -\cos t
\end{pmatrix} \begin{pmatrix}
  dN_1(T) \\
  dN_2(T)
\end{pmatrix}.
$$

(17)

To determine the $f_j$’s and $\Sigma_k$’s we equate (13) to (17), identifying the leading order $O(1)$ terms and the higher order corrections with coefficient $e^2$ or $g_j$. This results in a series of equations, from which we obtain the drift and diffusion coefficients. As noted above, we consider the parameter range in which $e^2 \ll 1$ and $g_j \ll 1$, as discussed further in Section 4.

The details of the calculation are given in the Appendix. Since the terms in (13) and (17) are functions of both fast and slow time, an essential element of the calculation is the projection of the equations onto the fast oscillatory modes in order to isolate the slow modulation over long time scales. This projection is identical to the standard solvability condition used in normal form calculations to eliminate secular (unphysical) terms which grow linearly in time [25, 26]. In deterministic systems the projection leads to amplitude equations on a slow time scale, equivalent to an averaging over the fast time scale. In the stochastic system the projection plays a similar role, producing stochastic amplitude equations with an appropriate
approximation for the noisy excursions on the slow time scale [18, 28]. It has the form
\[ \int_0^{2\pi} (u^*, v^*) \cdot (28) \, dt \] (18)
where (28) is the result from equating (13) and (17) and \((u^*, v^*)\) is the solution of the system adjoint to the deterministic part of (8). Under the multi-scale assumption, those functions of the slow time \(T\) in (18) are treated as constants in the integration.

Then, as shown in the Appendix, we find the drift coefficients in (14),
\[ f_1 = -A, \quad f_2 = -B. \] (19)

The coefficient matrices for the diffusion terms in (14) are
\[ \Sigma_1 = \frac{g_1}{2be} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \Sigma_2 = \frac{g_2}{2c} \begin{pmatrix} -b & 1 \\ -1 & -b \end{pmatrix}, \quad \Sigma_3 = \frac{g_2}{2eb} \begin{pmatrix} 0 & -b \\ b & 0 \end{pmatrix}. \] (20)

### 3.1 Identification of the Stationary Process \((A, B)'\)

The stochastic amplitude equations (14) for \(A\) and \(B\) are
\[ d\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} \, dT + \Gamma \, dW(T), \] (21)
where
\[ \Gamma = \frac{1}{2be} \begin{pmatrix} g_1 & -b^2g_2 & bg_2 & -bg_2 & 0 & 0 \\ g_0 & -bg_2 & -b^2g_2 & 0 & g_1 & bg_2 \end{pmatrix} \]
and \(dW(T)\) is the column vector \((dw_{11}, dw_{21}, dw_{22}, dw_{32}, dw_{12}, dw_{31})'\). We see from the form of (21), that \((A(T), B(T))'\) is an Ornstein-Uhlenbeck type process, and therefore we can write the stationary distribution for \(A\) and \(B\) explicitly. This distribution gives us important information about the parameter regions in which we expect to see coherence resonance. To see that the stationary process has mean \((0, 0)\) we begin by taking the expected value on both sides of equation (21). Solving the resulting equations we obtain the means of \(A\) and \(B\): \(E[A] = C_1 e^{-T}\) and \(E[B] = C_2 e^{-T}\) where \(A(0) = C_1\) and \(B(0) = C_2\) for constants \(C_1, C_2\). This implies that both \(E[A]\) and \(E[B]\) approach zero as \(T\) increases. The one-time-point stationary law of the Ornstein-Uhlenbeck process can be identified as a bivariate normal distribution, \(N(0, K)\) with mean zero and covariance matrix \(K\), using formulae in, for example, [29], where
\[ K = \text{cov}(A(T), B(T))' \]
\[ = f_0^\infty \exp \left[ \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} s \right] \Gamma \Gamma' \exp \left[ \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} s \right]' \, ds \]
\[ = \frac{1}{2eT} (g^2_1 + (1 + b^2) b^2 g^2_2 + b^2 g^2_2) f_0^\infty \exp \left[ 2 \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} s \right] ds \]
\[ = \frac{\delta^2}{2e} I f_0^\infty \exp[-2s] ds = \frac{\delta^2}{2e} I. \]

The \(2 \times 2\) identity matrix is represented by \(I\). We observe that \(A(T)\) and \(B(T)\) are independent, each with variance \(\delta^2/(2e^2)\). The complete distribution of the
stationary Gaussian process process \((A, B)\)' is then identified by its auto-covariance matrix at time points \(T\) and \(U\),

\[
\text{cov}[(A(T), B(T))', (A(U), B(U))'] = \frac{\delta^2}{2\varepsilon^2} e^{-|U-T|} \mathbf{I}.
\]  \hspace{1cm} (22)

We can replace (21) with a system of amplitude equations driven by two independent Brownian motions \(w_1\) and \(w_2\),

\[
d \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} dT + \frac{\delta}{\sqrt{2\varepsilon}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} dw_1 \\ dw_2 \end{pmatrix}.
\] \hspace{1cm} (23)

### 4 Parameter ranges for coherence resonance

We identify the parametric conditions under which one expects to observe the phenomenon of coherence resonance, and demonstrate that under these conditions the multi-scale construction of the previous section captures the character of the SIR process.

From the amplitude equations (23) we see that the ratio \(\delta/\sqrt{2\varepsilon}\) is a measure of the magnitude of the noise in the process \((A, B)\), with \(\delta^2/2\varepsilon^2\) as the variance of the stationary process as shown in (22). Hence one condition for the coherence resonance phenomenon is in terms of \(\delta^2/2\varepsilon^2\). For very small values of \(\delta^2/2\varepsilon^2\) one expects to see relatively small oscillations. For \(\delta^2/2\varepsilon^2\) larger but still well below unity, the stochastic fluctuations balance with the deterministic slow decay, so that both stochastic and deterministic features are apparent in the dynamics. For \(O(1)\) variance, the noise dominates the behavior, so that the PSD is no longer concentrated around a single frequency (Figure 3). Furthermore, the ratio \(\delta^2/2\varepsilon^2\) provides a criterion for the application of the multi-scale analysis: the method is valid for parameter ranges in which the dynamics is not dominated by the noise. If \(\delta^2/2\varepsilon^2\) is large, the stochastic variations govern the dynamics, so that an approximation based on a slowly varying modulation is no longer appropriate.

The validity of the multi-scale analysis can then be quantified through the two criteria:

\[
\varepsilon^2 \ll 1, \quad \frac{\delta^2}{2\varepsilon^2} \ll 1.
\] \hspace{1cm} (24)

The first of these criteria is the standard assumption in a multiple scales expansion, which states that the crucial time scales are well separated, while the second is the condition on the noise as described above. In terms of the biological parameters these criteria are:

\[
\varepsilon^2 = \frac{R_0}{2} \sqrt{\frac{\mu}{\mu + \gamma R_0} - 1} \ll 1, \quad \hspace{1cm} (25)
\]

\[
\frac{\delta^2}{2\varepsilon^2} = \frac{\mu + \gamma}{4N\mu} \left(1 + \frac{R_0 + 1}{R_0 - 1} + 2\frac{\mu + \gamma}{\mu(R_0 - 1)}\right) \ll 1. \hspace{1cm} (26)
\]

Figure 2 shows the parameter regions in the \(\gamma-R_0\) plane where these criteria are valid, for two different values of \(\mu\) and \(N\).

From (25) and Figure 2 it is clear that the boundary of the parameter region corresponding to \(\varepsilon^2 \ll 1\) does not depend on \(N\). We consider this criterion for \(R_0\)
both near and away from unity, since the factor $R_0 - 1$ plays a significant role in each. For small values of $\mu$, (25) can be written approximately as $\frac{R_0}{2} \sqrt{\frac{\gamma}{R_0 - \gamma}} \ll 1$. Then for $R_0$ away from 1, this criterion is satisfied for reasonably large values of $\gamma/\mu$. The range of validity for (25) is inside the dotted line and diamonds in Figure 2. We observe that for $\mu = 1/55$ this criterion is satisfied for a large region in the $\gamma$-$R_0$ plane. For increasing $\mu$ the condition is violated for moderate values of $\gamma$ and larger $R_0$.

Now we consider the second criterion (26), which restricts the variance of the stochastic fluctuations relative to the slow time scale. For small $\mu$, the approximation to (26), $\frac{\gamma^2}{2N\mu^2(R_0 - 1)} < 1$, captures this criterion over much of the $\gamma$-$R_0$ plane. The range of validity for (26) is below the o’s, solid, and dash-dotted lines in Figure 2. For $R_0$ approaching unity, the criterion is satisfied only for small $\gamma$. For $R_0$ away from unity, the size of the region decreases as $\mu$ decreases and increases with increasing $N$.

The range of parameters most favorable for sustained coherence resonance is where $R_0$ is away from unity, and $\gamma/\mu$ is large but bounded in the sense that it is still an order of magnitude smaller than $\sqrt{N}R_0$. The ratio $\gamma/\mu$ is interpreted as the ratio of an average human lifetime to duration of an infection in an individual. For most infectious diseases this ratio is indeed large. For large $N$ and $R_0$ away from unity this combination produces a relatively slow decay of the inherent oscillations ($\epsilon^2 \ll 1$), while the random effects are strong enough to sustain near periodicity without dominating the system or driving it into a qualitatively different behavior (large oscillations and/or extinction). Since $R_0 \approx \beta/\gamma$, the range $R_0 - 1 = O(1)$ suggests cases where $\beta$, the number of effective contacts per unit time, is at least double the recovery rate $\gamma$. In Figure 4 we show the typical parameters in the $R_0 - \gamma$ plane for a number of diseases [11] [30].

In Figure 3 we compare the PSD’s for the full system and the multi-scale approximation. In Figure 3a,d,e, and f the parameter values are well within the regions bounded by the criteria (25) and (26), and consequently there is good agreement between the PSD of the multi-scale approximation and the full system. The violation of the criterion (26) is manifested by an effective increase in the random fluctuations over the long time scale $T$, with the noise supporting larger excursions from the equilibrium. In this range there are still oscillations, but they do not appear regular since they are mainly noise driven. Increasing the fluctuations includes the nonlinear effects which exhibit other modes of oscillation, so that the PSD is no longer characterized by a single peak. Then the single mode approximation (12) used for the linearized problem is no longer sufficient to capture these higher order effects. In Figure 3b,c the parameter values are closer to the boundary for the criterion (26). In Figure 3b the small peak in the PSD at $2\Omega$ is more pronounced, together with other contributions beyond the dominant peak at $\Omega$. In Figure 2c, with even larger values of $\delta/\sqrt{2}\epsilon$, the PSD is not concentrated at a single frequency, indicating a noise driven process.

Figure 2f shows results for a larger value of $\mu = .1$, for the same parameter range used in the 2004 numerical studies of the persistence of influenza using SIR and the qualitatively similar SIRS model in [1] and [2]. The focus of [1] is on recurrent epidemics of influenza when immunity is “lost” due to a change in the dominant strain. Then members of the removed class return to the class of susceptibles, so
that the appropriate model is SIRS with a larger $\mu$ representing the return rate. The choice of parameters is further motivated by the interest of [1], [2] in the effect of seasonal forcing in a stochastic model with sustained oscillations where the natural period is about the same as the one year forcing ($\Omega \approx 6$). Indeed they observed oscillations of this type in stochastic simulations, and our analysis shows that this parameter combination also falls in the range where coherence resonance is expected.

In Figure 5 we compare the approximation of the stationary densities, characterizing the long time behavior of the infected population. In this study we have kept only the linear contributions to the multi-scale approximations, so it does not capture the variation in the tails of the distribution. Nevertheless, for cases where there is good agreement in the PSD, as shown in Figure 3, the linear multi-scale approximation captures the bulk of the density. In order to capture additional variation, nonlinear effects would have to be included. Extensions of the stochastic multi-scale analysis for nonlinear oscillations have been illustrated in [18]. These straightforward extensions could be applied to this model to capture more of the variance and tail behavior of the density, due to nonlinear interactions.

We also show the density for the choice of parameters corresponding to Figure 3c, where (26) is violated. As would be expected, the multi-scale approximation (shown by the dash-dotted line in Figure 5b) does not give a reasonable approximation to the dynamics. In this noise dominated case, the infective population can reach extinction over a short time period. In the simulation we allow random fluctuations to restart the disease and often these near-extinction periods are followed by large fluctuations in the infected population. In these cases one has to include additional complications in the model and analysis, and we leave these extensions for future study.

5 Discussion

A multiscale analysis is applied to approximate, in a wide parameter range, nearly regular oscillatory dynamics observed in stochastic SIR. The fluctuations can be captured by a combination of sinusoids where the amplitude coefficients are an Ornstein-Uhlenbeck process running on a slower time scale. The approximation illustrates the two main ingredients for this manifestation of coherence resonance: the slow time scale of the inherent oscillation’s decay combined with the resonance of the small noise level with this mode. The analysis provides a prediction of the parameter range where one would expect to observe this behavior, with the parameters of several common infectious diseases falling in this region. The approximation is based on the multiscale dynamics in a neighborhood of the endemic equilibrium point of the deterministic SIR which exists when the basic reproductive number of the process, $R_0$, is larger than 1. The nearly regular nature of the fluctuations, and the parameter regime in which it can be observed, is verified via a comparison of the power spectral densities of the stochastic SIR and the multi-scale approximation. We also compare the stationary distributions of the two processes.

The interest in this result stems from observed epidemic oscillations in, for example, measles and influenza. Whereas sustained oscillations can be produced in deterministic models by introducing various complications, we show here why they are to be expected in a wide range of diseases, resulting simply from the presence
of the coherence resonance between the random variations and damped oscillations. The multiscale method will produce a similar result for other stochastic disease models where the deterministic version shows damped oscillations to an endemic equilibrium. If in addition a periodic forcing is present with frequency near to the natural frequency corresponding to the model parameters, then the amplitude of the autonomous stochastic resonance will be greatly enhanced as indicated numerically already in [2] and [1]. The pattern of recurrent epidemics may be influenced by several additional factors, for example, stratified population, quarantine, multiple strains. These factors, added to the basic model, may provide additional sources for sustained oscillations. We expect that the multiscale approach presented here would be valuable for the study of more complicated models also, providing a means to analyze the different sources of oscillations.

The phenomenon of sustained oscillations occurs if an external stochastic element is present, such as arbitrary additive noise or parametric noise, instead of, or in addition to, the stochastic fluctuations about the averaged behavior given in (5). Furthermore, the oscillations with the same stationary character persist even though the initial state is the equilibrium. That is, they are inherent in the system and do not depend on the appearance of large epidemic spikes.

The multiscale method leads to explicit criteria (25) and (26) for coherence resonance, and thus provides an insight into the dynamics which would not be provided by a standard exact solution of the linear problem. The results are consistent with other studies, even where the asymptotic construction of the envelope has not been pursued. For example, in the study of McKane et al. [5] of a population model, one can nondimensionalize as in Section 1. The parameters chosen in that study correspond to \( \epsilon^2 \approx .15 \) and \( \delta^2/2\epsilon^2 \approx .1 \). For these values of the parameters agreement was observed in [5] between the PSD’s of the linearized stochastic system and that of the full system, as is predicted by the criteria (25)-(26) provided by the multiscale analysis of this paper. The stochastic multi-scale method has been applied in a number of other applications where coherence resonance is observed [6] [18]. These studies illustrate the generalizations of the approach in applications which have a number of additional factors: nonlinear effects, delays, phase synchronization and multiplicative noise.

6 Appendix

Equating (13) and (17) we separate terms into those which are \( O(1) \) on the fast time \( t \) scale, and those which are higher order: that is, those terms with coefficient \( O(\epsilon^2) \) or coefficient \( g_j \) for \( j = 1, 2 \), when the equations are written on the \( t \) time scale. Then the \( O(1) \) terms are

\[
\left( \begin{array}{c}
-bA \sin t + Bb \cos t \\
(A \cos t + B \sin t)
\end{array} \right) dt = \left( \begin{array}{c}
b(-A \sin t + B \cos t) \\
A \cos t + B \sin t
\end{array} \right) dt
\]

which is an identity following directly from the fact that the proposed form of solution (12), treating \( A \) and \( B \) as constants, solves the linearized deterministic problem to leading order. Following the cancellation on the \( O(1) \) scale, we are left
with the higher order corrections for $\epsilon^2 \ll 1$ and $g_j \ll 1$,

\[
\begin{pmatrix}
  b \cos t f_1 + b \sin t f_2 \\
  \sin t f_1 - \cos t f_2
\end{pmatrix} dT + \begin{pmatrix}
  b \cos t \\
  \sin t
\end{pmatrix}
\begin{pmatrix}
  b \sin t \\
  -\cos t
\end{pmatrix}
\begin{pmatrix}
  N_1(T) \\
  N_2(T)
\end{pmatrix} = \\
\begin{pmatrix}
  -2b[A \cos t + B \sin t] \\
  0
\end{pmatrix} dT + \begin{pmatrix}
  g_1 dW_1 - b^2 g_2 dW_2 \\
  g_2 dW_2 - g_2 dW_3
\end{pmatrix},
\]  

(28)

where we have used the fact that $dT = \epsilon^2 dt$.

Note that the expressions on both sides of (28) have terms on both the fast time scale $t$ and the slow time scale $T = \epsilon^2 t$. Using the method of multiple scales, we look for a projection of these equations on the fast time scale $t$, integrating over one period of the fast oscillations. The resulting expressions vary on the slow time scale $T$ only, yielding equations for the coefficients $A$ and $B$ in (14).

In particular, we project (28) onto the solution $(u^*, v^*)$ of the noise-free linear system adjoint to (12),

\[
\begin{pmatrix}
  u^*_1 \\
  v^*_1
\end{pmatrix} = \begin{pmatrix}
  \cos t \\
  b \sin t
\end{pmatrix}, \quad \begin{pmatrix}
  u^*_2 \\
  v^*_2
\end{pmatrix} = \begin{pmatrix}
  -\sin t \\
  b \cos t
\end{pmatrix}.
\]  

(29)

Under the multiscale assumption the functions of the slow time, $T$, are treated as independent of the fast time $t$. Without loss of generality we compute the projections treating the drift and diffusion coefficients separately.

First we consider the drift terms in (28),

\[
\begin{align*}
b \cos t f_1 dT + b \sin t f_2 dT &= -2b(A \cos t + B \sin t) dT, \\
-cos t f_2 dT + sin t f_1 dT &= 0,
\end{align*}
\]  

(30)

and we project (30) onto the solution to the adjoint of (12)

\[
\int_0^{2\pi} [u^*_j(b \cos t f_1 + b \sin t f_2) + v^*_j(- \cos t f_2 dT + \sin t f_1)] dt = \\
\int_0^{2\pi} u^*_j(-2b(A \cos t + B \sin t)) dt,
\]  

(31)

for $j = 1, 2$, where we have integrated over one period of the fast oscillation. Using the multi-scale assumption, we integrate in (31), treating functions of $T$ as independent of $t$. The resulting equations for the drift coefficients in (14) are

\[
2bf_1 = -2Ab \quad -2bf_2 = 2Bb,
\]  

(32)

which yields (19).

Now we consider the diffusion coefficients in (28). Using well-known properties of Brownian motion, we rewrite $g_j dW_k(t)$ as

\[
g_j dW_k(t) = \frac{g_j}{\epsilon} [\cos t dw_{k1}(T) + \sin t dw_{k2}(T)]
\]  

(33)

where $w_{kj}$ are independent standard Brownian motions. Substituting (33) into (28), restricting our attention to the noise terms, and projecting as in (31), we obtain the
system

\[
\int_0^{2\pi} u_j^*(\frac{g_1}{\epsilon} [\cos tw_{11}(T) + \sin tw_{12}(T)] - b^2 \frac{g_2}{\epsilon} [\cos tw_{21}(T) + \sin tw_{22}(T)])
\]

\[
+ v_j^*(\frac{g_2}{\epsilon} [\cos tw_{21}(T) + \sin tw_{22}(T)] - \frac{g_2}{\epsilon} [\cos tw_{31}(T) + \sin tw_{32}(T)])dt =
\]

\[
\int_0^{2\pi} u_j^*(b \cos tN_1 + b \sin tN_2) + v_j^*(\sin tN_1 - \cos tN_2)dt,
\]

(34)

for \( j = 1, 2 \). Now we substitute (29) into (34) and integrate with respect to \( t \), treating functions of \( T \) as independent of \( t \), to obtain

\[
g_1 dw_{11} - b^2 g_2 dw_{21} + bg_2 dw_{22} - bg_2 dw_{32} = 2beN_1,
\]

\[
-g_1 dw_{12} + b^2 g_2 dw_{22} + bg_2 dw_{21} - bg_2 dw_{31} = -2beN_2.
\]

\[\Rightarrow 2be \left[ \Sigma_1 \left( \begin{array}{c} d\xi_{11}(T) \\
\frac{d\xi_{12}}{dT}(T) \end{array} \right) + \Sigma_2 \left( \begin{array}{c} d\xi_{21}(T) \\
\frac{d\xi_{22}}{dT}(T) \end{array} \right) + \Sigma_3 \left( \begin{array}{c} d\xi_{31}(T) \\
\frac{d\xi_{32}}{dT}(T) \end{array} \right) \right] = (35)
\]

We are free to identify each \( d\xi_{ij} \) with \( dw_{ij} \) to obtain the coefficient matrices (20).

Note that in the application of the projection in (34), we have treated the Brownian motions \( w_{kj}(T) \) as independent of \( t \). This assumption appears to be suspect, since Brownian motion has variation on all time scales. However, all oscillatory modes except for (12) decay strongly on the fast \( t \) scale, so that the treatment of \( w_{kj}(T) \) as independent of \( t \) is an approximation which neglects those variations which rapidly decay.

Acknowledgements: We thank the Sloan Foundation, NSF Grant DMS-0441114, NSA Grant H98230-05-1-0097, and NSERC Discovery Grant for support. We are also grateful to Carlos Castillo-Chavez for several helpful discussions.

References


Figure 2: Parameter regions in the $R_0$-$\gamma$ plane corresponding to the two criteria (25)-(26). (Left) For $\mu = 1/55$, the boundary $\frac{\delta^2}{2\sigma^2} \leq .2$ is shown for $N = 500,000$ (dotted line) and $N = 2,000,000$ (solid line). The boundary for $\epsilon^2 \leq .1$ (diamonds) is the same for both $N = 500,000$ and $N = 2,000,000$. The region where multi-scale coherence resonance is expected is inside the V-shaped curve ($\epsilon^2 < .1$) and below the dotted or solid line for $N = 500,000$ or $N = 2,000,000$ ($\frac{\delta^2}{2\sigma^2} \leq .2$), respectively. (Right) For $N = 500,000$ the boundary $\frac{\delta^2}{2\sigma^2} \leq .2$ is shown for $\mu = 1/55$ (dotted line) and $\mu = .1$ (dash-dotted line). The boundary for $\epsilon^2 \leq .1$ is shown for $\mu = 1/55$ (diamonds) and $\mu = .1$ (circles). The region where multi-scale behavior is expected is both inside the V-shaped curve ($\epsilon^2 < .1$) given by diamonds or circles for $\mu = 1/55$ or $\mu = .1$, respectively, and below/to the right of the dotted or dash-dotted ($\frac{\delta^2}{2\sigma^2} < .2$) for $\mu = 1/55$ or $\mu = .1$, respectively. For increasing $\mu$, the multi-scale region in the $R_0$-$\gamma$ plane increases in size and and shifts to larger values of $\gamma$. For increasing $N$, the multiscale behavior is expected over a larger range of $\gamma$. The *’s are located at the parameter values used in the PSD plots in Figure 3, with the letter labels on the *’s matching the corresponding subplots shown in Figure 3. We see that letters a,d,e, and f are in the coherence resonance region, and the corresponding PSD’s in Figure 3 have strong peaks at the predicted frequency. Case b is close the boundary for this behavior and exhibits some mixed dynamics, while case c is outside the coherence resonance region and consequently shows irregular oscillations, as confirmed by the corresponding PSD.
Figure 3: Graphs of the PSD (vertical axis) vs. frequency of the oscillations, comparing the PSD’s for the full stochastic model (5) (dash-dotted line) and the multi-scale approximation (12) (solid line). The parameters for the subplot are:

a) \( R_0 = 15, \gamma = 20, \mu = 1/55, N = 500,000; \)
b) \( R_0 = 15, \gamma = 25, \mu = 1/55, N = 500,000; \)
c) \( R_0 = 15, \gamma = 30, \mu = 1/55, N = 500,000; \)
d) \( R_0 = 15, \gamma = 30, \mu = 1/55, N = 2,000,000; \)
e) \( R_0 = 7, \gamma = 33, \mu = 1/55, N = 2,000,000; \)
f) \( R_0 = 10, \gamma = 35, \mu = 1/10, N = 500,000; \)

which are compared to the regions for stochastic multi-scale behavior in Figure 2. In cases a,d,e,and f the PSD has a strong peak at the deterministic frequency, consistent with the parametric conditions for coherence resonance. The horizontal axis for each is in terms of the original variables, so that the peak is observed at the deterministic frequency \( \Omega = \sqrt{\frac{3\mu}{R_0}}(R_0 - 1) \). Case b is close the boundary for coherence resonance, so there are some larger contributions from other frequencies, while case c is outside the coherence resonance region, so that its PSD is not sharply peaked and the multi-scale approximation is not valid.
Figure 4: Stars indicate typical parameters for $R_0$ and $\gamma$ for some childhood diseases. The boundaries for the criteria (25) and (26) are shown for $\mu = 1/55$ and $N = 2000000$ as in Figure 2.

Figure 5: Comparison of the numerical computation for the stationary density $p(I)$ for the infected population $I$, using the multiscale approximation (solid and dash-dotted lines) and the full system (‘s and diamonds). In all cases the density is approximated using 2000 realizations for a value of $t > 100$. On the left, we show the density for the same parameters as in Figure 3a (‘s and dash-dotted line) and Figure 3f (diamonds and solid line). On the right, we show the density for the same parameters as in Figure 3c (‘s and dash-dotted line), where the parameter values violate (26), and Figure 3d (diamonds and solid line), where the criterion (26) is satisfied.