On Identification of Continuous Time

Stochastic Processes

Jeremy Berkowitz
Federal Reserve Board

December 28, 1999

Address correspondence to:

Jeremy Berkowitz
Trading Risk Analysis
Federal Reserve Board
20th and C Streets, N.W.
Washington, D.C.
jberkowitz@frb.gov

Abstract: In this note we delineate conditions under which continuous time stochastic processes can be identified from discrete data. Such identification lies at the heart of the leading models of asset price and interest rate dynamics. The identification problem is approached in a novel way. The distribution of the observed stochastic process is expressed as the underlying true distribution, f, transformed by some operator, T. Using a generalization of the Taylor series expansion, the transformed function T∘f can often be expressed as a linear combination of the original function f. By combining the information across a large number of such transformations, the original measurable function of interest can be recovered.

Keywords: Identification, Aliasing

1. INTRODUCTION

It is a well-known fact that continuous time stochastic processes cannot, in general, be inferred from (evenly-spaced) discretely sampled data. In recent years, however, there has been a resurgence of interest in this class of identification problems. Such research has been spurred by a desire to estimate stochastic differential equations describing the dynamics of asset prices and interest rates (e.g., Lo (1988), Ait-Sahalia (1996a,b), Anderson and Lund (1997), Stanton (1997)). Theoretical developments, such as Hansen and Scheinkman's (1995) extension of the Generalized Method of Moments to continuous time models have had to rely on strong assumptions in order to deal with the aliasing problem. Duffie and Singleton (1993) and Gallant and Tauchen (1996) devise simulation-based estimation methods with similar stipulations.

The vast majority of models of asset price dynamics, models of the term structure of interest rates and derivative pricing formulae are based on the assumption of an underlying continuous time process. However, conclusions regarding the empirical relevance of such models are reached via estimation and evaluation methods which use discrete data. The relationship between the theoretical model and estimated model is thus of paramount importance in interpreting empirical results.

In the present paper, we approach the identification problem in a novel way. The distribution of the observed stochastic process is expressed as the underlying true distribution, f, transformed by some operator, T. Using a generalization of the Taylor series expansion, the transformed function T∘f can often be expressed as a linear combination of the original function f. By combining the information across a large number of such transformations, the original measurable function of interest can be recovered.

This approach substantially widens the class of identifiable models relative to Hansen and Sargent (1983), Hansen and Scheinkman (1995), and Hansen, Scheinkman and Touzi (1998). For example, in the context of diffusion models, we show that the infinitesimal generator that characterizes local dynamic behavior can be arbitrarily well approximated by linear combinations of conditional expectations operators.

An alternative way to view our approach is that we combine the information corresponding to different sampling rates to approximate the true model arbitrarily well. Section

2

2 describes the basic mathematical framework. Section 3 presents some illustrative stochastic processes of interest. Section 4 concludes.

2. THE EXPONENTIAL FORMULA

Consider a semigroup of operators indexed by h, T_h, from $C[0,\infty)$ to $C[0,\infty)$. Assume that T_h is continuous in the sense that $T_0 = I$ and $T_h \to I$, as $h \to 0$ (from the right), where I is the identity operator.

A remarkable result due to Hille and Phillips (1957) is the generalized Taylor formula,

$$T_h \circ f = \sum_{r=0}^{\infty} h^r / r! \ A^r \circ f \tag{1}$$

where f is any bounded continuous function on $[0,\infty)$, $A = \lim_{h \to 0^+} h^{-1}(T_h - I)$ is the infinitesimal generator and $A^r \circ f$ is r-th iterated application of A to f. As usual, the notation A^0 should be taken to mean the identity operator. The convergence in (1) is uniform in h. This expansion, also referred to as the exponential formula of semi group theory, does not require differentiability.[1] Moreover, it defines expansions of not just functions but very general families of *operators.*

The importance of the exponential formula is that it expresses a transformation of a given measurable function as a linear combination of the original untransformed function. The particular measurable functions of interest to us may be random variables, their distributions or other statistics of interest. In many circumstances such as temporal aggregation, the econometrician can only observe some transformation of the random variable(s) of interest.

A leading example is estimation of stochastic differential equations describing the dynamics of interest rates (e.g., Lo (1988), Anderson and Lund (1997), Ait-Sahalia (1996a,b) and Stanton (1997)). Since continuous data is not available, discretization (aggregation)

---

[1]It is easy to see why the formula bears this alternative name. The matrix exponential of a matrix A is defined as

$$e^A = \sum_{m=0}^{\infty} \frac{A^m}{m!}$$

by, for example, Coddington and Levinson (1955, p.64).

complicates classical estimation procedures. The transition densities associated with discrete observations depend on the drift and diffusion functions in generally unknown ways. The underlying stochastic process cannot generally be inferred from the conditional expectations operators associated with discrete data.

The utility of equation (1) is that in many cases deformations of the underlying process can be expressed as linear functions of the true underlying process. The expansion cannot be "inverted" in the sense that exp(Af) does not generally identify Af. However, in this article we discuss a method of combining the information in sets of expansions which is sufficient to approximate Af arbitrarily well.

In particular, we can use the generalized Taylor series to write,

$$
\begin{aligned}
T_h \circ f &= f + hA \circ f + [hA]^2 \circ f \,/2 \;+... \\
T_{2h} \circ f &= f + 2hA \circ f + [2hA]^2 \circ f \,/2 \;+...
\end{aligned}
$$

$$...$$

(2)

for any multiple of h. Inspection of the expansions in (2) suggests that it might be possible to combine the information across expansions to recover the leading term, f.

In the next proposition, we show that this is indeed the case. Moreover, it will become apparent that any desired term in the expansion can be recovered by similar methods. It will turn out that in many cases the second term of the expansion, Af, will be of particular interest. The infinitesimal generator, when it exists, describes the local behavior of stochastic processes.

As an illustration, consider the first order linear combination, $2T_h \circ f - T_{2h} \circ f = f - h^2 A^2 \circ f$. This combination yields an estimate of f with error of size $h^2 A^2 \circ f$. We would like to establish a method of approximating f such that the error term can be made arbitrarily small.

Since it should cause no confusion, we will use operator notation and suppress dependence on the test function, f. For example, we would write (1) more compactly as $T_h = \sum h^r/r! \, A^r$.

Before proceeding, we will need to make following assumption.

ASSUMPTION A1: The operator T_h is a weak contraction. That is, $\|T_h \circ f\| \le \|f\|$ for all $f \in \mathcal{L}^2$, the space of square integrable measurable functions. In operator notation, $\|T_h\| \le 1$ by which we mean that $\|T \circ f\| \le \alpha \|f\|$ for $\alpha = 1$.

The contraction property is not overly restrictive in probabilistic contexts. For example, consider convolution operators

$$T_h f(x) = \int_{-\infty}^{\infty} f(x-y) \ dF_h(y)$$

where $F_h(y)$ is a given family of distribution functions indexed by h. It is known that convolution operators are weak contractions (e.g., Feller (1971), p. 257). For our purposes, leading examples include the expectations operator and the translation operator.

Without loss of generality let h=1. Let w(k) be the k-vector that solves the linear equations implied by k of the expansion in equation (2),

$$w(k)' \begin{bmatrix} 1 & 1 & 1/2! & ... & 1/k! \\ 1 & 2 & 2^2/2! & ... & 2^k/k! \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & k & k^2/2! & ... & k^k/k! \end{bmatrix} = [1 \ 0...0]$$

It should be clear that w(k) depends on the order, k, so we will suppress this dependence for notational simplicity. By construction, the linear combination $\sum_{i=1}^{k} w_i T_i = I + n(k) A^k/k!$ where $n(k) = \sum_{i=1}^{k} w_i i^k$.[2] This leads us to the following identification result.

PROPOSITION 1: Under assumption A1, the function f can be approximated to within any desired

---

[2]The weights, w_i, can also be calculated by combinatoric methods. In particular, $w_i(k)$ is given by $\binom{k}{i}(-1)^{i+1}$, a formula that arises naturally in the context of iterated bootstrap bias correction (e.g., Hall (1992)).

accuracy by linear combinations of $\{T_h:h>0\}$ in the sense that $\sum_{i=1}^{k} w_i T_i f \rightarrow f$ in the L^2 norm.

PROOF:

As indicated, the linear combination $\sum_{i=1}^{k} w_i T_i \rightarrow I + n(k)/k! \ A^k$. We will proceed by showing that $|n(k)/k!| = O(1)$ and that $\|A^k\| \rightarrow 0$. Combined these two results, the Cauchy-Schwartz inequality implies that $\|n(k)/k! \ A^k\| \leq \|n(k)/k!\| \ \|A^k\| \rightarrow 0$.

First consider the asymptotic behavior of $n(k)/k!$ which is given by

$$\sum_{i=1}^{k} w_i \frac{i^k}{k!}$$

$$= \sum_{i=1}^{k} w_i \frac{i^k}{k^k} \frac{k^k}{k!}$$

$$\approx \sum_{i=1}^{k} w_i \left(\frac{i}{k}\right)^k e^k$$

$$= \sum_{i=1}^{k} w_i \left(1 - \frac{k-i}{k}\right)^k e^k$$

where we have used the fact that $k^k/k!$ is approximately $\exp(k)$ for large k. It is easy to show that as $k \rightarrow \infty$, $w_i(k) \rightarrow 0, \forall \ i<k$ but equals ± 1 for $i \equiv k$ so that $n(k) = \pm 1$.

Now it remains to consider $\lim \|A^k\|$. It is straightforward to show that the weak contraction property of T_h implies that A is itself a contraction. In particular, for any $h>0$ $\|\exp[h^{-1}(T_h \circ f - f)]\| = \|\exp[h^{-1}T_h \circ f - h^{-1}f]\| \leq \exp\|h^{-1}T_h \circ f - h^{-1}f\| \leq \exp(I)$ by the contraction property (see discussion in Feller (1971), p.353). Thus $\|A\|^k \rightarrow 0$ as $k \rightarrow \infty$ and the result follows by Hölder's inequality since $\|A^k\| \leq \|A\|^k \rightarrow 0$. ∎

Proposition 1 establishes conditions under which the leading term in the exponential formula can be recovered. Consider next approximation of the second term in the Taylor expansions, $A \circ f$. Let $v(k)$ be the k-vector that solves the linear equations,

$$v(k)^{/} \begin{bmatrix} 1 & 1 & 1/2! & ... & 1/k! \\ 1 & 2 & 2^2/2! & ... & 2^k/k! \\ : & : & : & : & : \\ 1 & k & k^2/2! & ... & k^k/k! \end{bmatrix} = [0 \ 1 \ 0...0]$$

By construction we have that $\sum_{i=1}^{k} w_i T_i = A + n(k)/k! A^k$.

PROPOSITION 2: The generator A can be approximated to within any desired accuracy by linear combinations of $\{\Im_t : t > 0\}$ in the sense that $\sum_{i=1}^{k} w_i \Im_i \rightarrow A$ in the uniform norm induced by the L^2 norm.

PROOF:

  By slight modification of the proof to proposition 1.■

  The implication of this result is that the infinitesimal generator can be approximated arbitrarily well with discrete observations.

3. EXAMPLES

  In this section, we illustrate the identification scheme in the context of some well known classes of continuous-time stochastic processes.

Diffusion Processes

  Consider an n-dimensional, stationary, continuous-time Markov process, $\{x_t\}$, defined on the probability triple $(\Omega, \mathfrak{Z}, \wp)$. Associated with this process is the semi-group of conditional expectation operators, $\{\Im_t : t \geq 0\}$, defined by

$$\Im_t f(y) = E[f(x_t)|x_0 = y]$$

where $f \in \mathcal{L}^2(Q)$, where Q is the stationary distribution of x_t.

  Let A be the associated generator,

$$Af(y) = \lim_{t \to o} \frac{1}{t} \ (\Im_t f - f) \tag{8}$$

which is well defined on a set of test functions, f (the limit is taken in the sense of the L^2-norm). Let D denote the subset of functions in L^2 for which (8) holds. As noted in Hansen and Scheinkman (1995), $\Im = \lim_{n \to \infty} (I - A/n)^{-n} = \exp(A)$. More generally, the exponential formula applied to the conditional expectation operator yields,

$$\Im_t \circ f = \sum_{r=0}^{\infty} \ (tA)^r / r! \circ f \ . \tag{9}$$

The exponential series $\sum_{k=0}^{\infty} T^k / k!$ is absolutely convergent in the uniform norm for every operator T with finite norm (e.g., Hirsch and Smale (1974), p. 83).

In this case, the first term in the exponential formula represents the stationary distribution. The second term, containing the infinitesimal generator of the semigroup of conditional expectations operators, describes the local behavior of the transitions. Recovery of the generator follows from proposition 2.

Aliasing in Absolutely Summable Processes

Assume that x_t is stationary with finite second moments. By the Wiener-Khintchine Theorem, the normalized integrated spectrum (the spectral distribution), $F(\omega)$, exists and is nonnegative, nondecreasing, and right continuous (Priestley (1981), p.219).

DEFINITION: Define the folding operator from $C[0,\infty]$ to $C[0,\infty]$, T_h, such that

$$T_h g(\omega) = \sum_{k=-\infty}^{\infty} \ g(\omega + 2\pi kh) \tag{10}$$

This operator is associated with wrapping the x-axis around a circle of length $2\pi h$. Its most familiar application relates the spectral density of a continuous time process to its discretely

sampled counterpart. That is, the spectral density of the process x_t sampled at frequency 1/h, is given by $T_h\, f(\omega)$ which we denote $f_h(\omega)$ for notational simplicity. The fact that $\hat{f}(\omega) = T_h \circ f(\omega)$ immediately indicates that $f(\omega)$ cannot, in general, be recovered from discrete data. This is the traditional formulation of the aliasing problem.

Rather than work with the spectral density, it will be convenient to work with the more general formulation:

$$\hat{F}(\omega) = \sum_{k=-\infty}^{\infty} F(\omega+2\pi kh) \tag{11}$$

where $F(\omega)$ is the spectral distribution function. Assumption B1 is sufficient to ensure the existence of $F(\omega)$.

ASSUMPTION B1: The process x_t possesses an absolutely summable autocorrelation function,
$\int_{-\infty}^{\infty} |\rho(\tau)|\ d\tau <\infty$.

This assumption is equivalent to assuming an absolutely continuous spectral distribution, ensuring the existence of $f(\omega)$, the spectral density (e.g., Priestley (1981), p.219).

Before moving to the main results of interest, it will be useful to establish two lemmas regarding the behavior of the spectral density. The proofs are given in the appendix.

LEMMA 1: As $\omega\to\infty$, $f(\omega)\to 0$.

Lemma 1 indicates that the spectrum asymptotes to zero. This result is not surprising in light of the fact that the integral of the spectrum, the variance, is assumed to be finite. Lemma 1 is used to establish the following result, a slight modification of the Integral Test.

LEMMA 2: Let $g(x)$ be a positive function defined on $[0,\infty)$ such that $\lim_{x\to\infty} g(x)=0$. Then $\sum_{x=1}^{\infty} g(x)$ converges if $\int_{0}^{\infty} g(x)dx$ converges.

In what follows, we establish the existence of the generator of $F_h(\omega)$ with respect to h. It will then be possible to apply the generalized Taylor series expansion as above.

PROPOSITION 3: Under assumptions B1, $AF(\omega) = \lim_{h \to 0^+} h^{-1}(F_h(\omega) - F(\omega))$ exists and is finite.
PROOF: See appendix.

Proposition 3 together with assumption B1 is sufficient to establish that the Taylor series expansion is valid and converges for all h. This allows $F(\omega)$ to be recovered arbitrarily well without continuously recorded data.

PROPOSITION 4: Under assumptions B1, $F(\omega)$ can be identified from data sampled at intervals of length $h \le 1$.
PROOF: See appendix.

Proposition 4 indicates that the aliasing problem disappears given suitable data availability. Put another way, the spectrum of any such continuous time model can be estimated with discretely-sampled data.

We do not require increasingly frequently sampled data, as for example, Bai and Perron (1998) do in the context of structural breaks. Of course, in any given application it is possible that the sampling frequency needed for consistent estimation is not available. Nevertheless, our result indicates that in a sense aliasing is not an inherent theoretical feature of continuous parameter models.

Possibly Nonstationary Linear Markov Models

The identification of continuous time stochastic processes has traditionally proceeded under the assumption of strict stationarity (e.g., Hansen and Sargent (1983), Hansen, and Scheinkman (1995), Ait-Sahalia (1996a) Hansen, Scheinkman and Touzi (1998)). In many circumstances, however, confining attention to stationary models would be quite restrictive. In this section we study loosening stationarity at the cost of narrowing the class of permissible models. In particular, replace assumption B1 with the following condition.

ASSUMPTION B2: x_t is described by a finite-dimensional first-order Markov model:

$$Dx_t = \Gamma x_t + \varepsilon_t \tag{12}$$

where D is the differential operator, ε_t is continuous time vector white noise with intensity matrix V_0. The square matrix Γ is real and finite.[3]

Discretely sampled data will have the form,

$$x_t = B_0 x_{t-1} + \tilde{\varepsilon}_t \tag{13}$$

where $B_0 = \exp(\Gamma)$ and $\tilde{\varepsilon}_t$ has intensity matrix $W_0 = \int_0^1 \exp(\Gamma\tau)V_0\exp(\Gamma'\tau)d\tau$.

Although we cannot use the methods based on spectral densities, the parameters governing the dynamics of the data, Γ, can be recovered in an analogous way. This is possible since the relation $B_0 = \exp(\Gamma)$ is obviously a manifestation of the exponential formula we have been exploiting throughout. In fact, a more general statement of the aggregation problem is that $B_0 = \Sigma (\Delta \Gamma)^n/n!$ where Δ is the sampling frequency. Using the information contained in multiple sampling frequencies it is possible to identify Γ.

PROPOSITION 5: Under assumption B2, Γ can be identified from data sampled at intervals of length $\Delta \le 1$.

PROOF:

$$
\begin{aligned}
B_\Delta - I &= \Delta\Gamma &+ 1/2\Gamma\Delta^2 &\quad + O(1/3!\Delta^3) \\
B_{2\Delta} - I &= 2\Delta\Gamma &+ 1/2\Gamma4\Delta^2 &\quad + O(2^3/3!\Delta^3) \\
B_{4\Delta} - I &= 4\Delta\Gamma &+ 1/2\Gamma16\Delta^2 &\quad + O(4^3/3!\Delta^3) \\
& & \cdots &
\end{aligned}
\tag{14}
$$

Taking the appropriate linear combination, produces $\Gamma + O(n(k)/k! \Delta^k)$. The remainder vanishes as $k \to \infty$ in parallel with the proof to proposition 4. ■

---

[3]Restricting attention to first-order Markov models is not consequential. Higher order models can be transformed into an expanded state-space such that the expanded process is Markovian (e.g., Priestley (1981)).

4. CONCLUSIONS

In this paper we deal with the identification problem by studying the information content in sets of discretely sampled data with increasingly low sampling frequencies. It had previously been established that first-order Markov models could be identified with restrictions on the intensity matrix of the innovation process (Hansen and Sargent (1983). Analogous results in Hansen and Scheinkman (1995) establish identification of diffusion processes which possess self-adjoint generators. In this paper, we show that such rotational restrictions are unnecessary.

It seems likely that the basic identification principle advocated here might be useful in contexts other than temporal aggregation. The application of generalized Taylor expansions may prove meaningful in more general identification and estimation problems in which an unwanted data transformation can be expressed as an operator.

REFERENCES

Ait-Sahalia, Y. (1996). "Nonparametric Pricing of Interest Rate Derivative Securities,"
Econometrica, 64, 527-560.

Ait-Sahalia, Y. (1996). "Testing Continuous-Time Models of the Spot Interest Rate," *Review of Financial Studies*, 9, 385-426.

Anderson, T. G. and J. Lund (1997), "Estimating Continuous-Time Stochastic Volatility Models of the Short-Term Interest Rate," *Journal of Econometrics*, 77, 343-377.

Bai, J. and Perron, P. (1998). "Testing for and Estimation of Multiple Structural Changes,"
Econometrica, 66, 47-79.

Coddington, E.A. and N. Levinson (1955). *Theory of Ordinary Differential Equations*. New York: McGraw-Hill.

Duffie, D. and K. Singleton (1993). "Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica*, 61, 929-952.

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*. New York: John Wiley.

Gallant, A. R. and G. Tauchen (1996). "Which Moments to Match?" *Econometric Theory*, 12, 657-681.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. New York: Springer-Verlag.

Hansen, L. P. and T. J. Sargent (1983). "The Dimensionality of the Aliasing Problem in Models with Rational Spectral Densities," *Econometrica*, 51, 377-387.

Hansen, L. P. and J. A. Scheinkman (1995). "Back to the Future: Generating Moment Implications for Continuous Time Markov Processes," *Econometrica*, 63, 767-804.

13

Hansen, L. P., Scheinkman, J. A. and Touzi, N (1998). "Spectral Methods for Identifying Scalar Diffusions," *Journal of Econometrics*, 86, 1-32.

Hille, E. and Phillips, R. S. (1957). *Functional Analysis and Semi-groups*. Publications of the American Mathematical Society.

Hirsch, M. W. and S. Smale (1974). *Differential Equations, Dynamical Systems, and Linear Algebra*. San Diego: Academic Press.

Lo, A. W. (1988), "Maximum Likelihood Estimation of Generalized Ito Processes with Discretely Sampled Data," *Econometric Theory*, 4, 231-247.

Phillips, P. C. B. (1973). "The Problem of Identification in Finite Parameter Continuous Time Models," *Journal of Econometrics*, 1, 351-362.

Priestley, M. B. (1981). *Spectral Analysis and Time Series*. London: Academic Press.

Stanton, R. (1997). "A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk," *Journal of Finance*, 52, 1973-2002.

Proof of Lemma 1

A necessary condition for convergence of any series, $\sum_0^\infty a_n$, is that $a_n \to 0$. Let $a_n = \int_{n\varepsilon}^{(n+1)\varepsilon} f(\omega)d\omega$

where $\varepsilon > 0$ so that $\sum_0^\infty a_n = \int_0^\infty f(\omega)d\omega$. It follows that a necessary condition for $\int_0^\infty f(\omega)d\omega$ to exist

(which must by assumption 2) is that $\int_{n\varepsilon}^{(n+1)\varepsilon} f(\omega)d\omega \to 0$ for all $\varepsilon > 0$. Since $\left| \int_{n\varepsilon}^{(n+1)\varepsilon} f(\omega)d\omega \right| > |f(n\varepsilon)|$, the

result follows. ∎

Proof of Lemma 2

Note that $\int_1^{n+1} g(x)dx = \sum_{k=1}^{n+1} \int_k^{k+1} g(x)dx \geq \sum_{k=1}^{n+1} g(k)$ where the inequality follows from the nonnegativity

of $g(x)$. ∎

Proof of Proposition 1

Write $F_h(\omega) - F(\omega) = \sum_{k=-\infty}^{\infty} F(\omega + 2\pi kh) - F(\omega)$. So

$AF(\omega) = 2\pi \sum kf(\omega + 2\pi kh)$, where $f(\omega)$ is the spectral density.[4]

  Since $\sigma^2 < \infty$, $\sum f(\omega + 2\pi kh)$ converges by Lemma 2. It remains only to show the

convergence of the infinite sum, $2\pi \sum kf(\omega + 2\pi kh)$. We will do this by application of the Root

Test. Namely, we show that $\limsup \sqrt[k]{kf(\omega + 2\pi kh)} < 1$.

  Suppose not, so that $\limsup \sqrt[k]{kf(\omega + 2\pi kh)} \geq 1$. Then $kf(\omega + 2\pi kh)$ would be greater than or

equal to 1 infinitely often and $f(\omega + 2\pi kh) \geq 1/k$ infinitely often. But this contradicts the fact that

$\sum f(\omega + 2\pi kh)$ converges. ∎

---

[4]Note that the derivative from the right always exists by Theorem 4.51 in Apostol (1974).

Proof of Proposition 4

Write the folding operator as expansions around the spectral distribution,

$$T_h F(\omega) = F(\omega) + A \circ F(\omega)h + 1/2 A^2 \circ F(\omega)h^2 + O(h^3)$$
$$T_{2h} F(\omega) = F(\omega) + A \circ F(\omega)2h + 1/2 A^2 \circ F(\omega)4h^2 + O((2h)^3) \qquad \textbf{(15)}$$
$$T_{4h} F(\omega) = F(\omega) + A \circ F(\omega)4h + 1/2 A^2 \circ F(\omega)16h^2 + O((4h)^3)$$
$$...$$

As before, appropriate linear combinations yield $F(\omega)$ plus a remainder term of $n(k)/k! \, h^k$. From the proof to proposition 1, $|n(k)/k!| = 1$. For $h < 1$ $h^k = o(1)$ and the result follows. For h=1, the remainder term is known, since $h^k \equiv 1$. ∎