Digest Journal of Nanomaterials and Biostructures Vol. 8, No. 3, July - September 2013, p. 1051 - 1059

STOCHASTIC APPROACHES IN PROTEIN SYNTHESIS-DEGRADATION - THE TWO INTERACTING PROTEIN CASE

D. DAVID-RUS*
National Institut for Physics and Nuclear Engineering (IFIN-HH), R-77125, Bucharest-Magurele, Romania

In this work, we study a protein synthesis degradation process by defining a general mathematical model. Using generating function technique we present a method that allows exact calculation of joint probability distribution of protein copies in a cell for a two dimensional birth-death process with interaction. We discuss the model in steady state for a particular choice of transition rules and fiind exact solutions.

(Received May 8, 2013; Accepted August 2, 2013)

Keywords: Protein synthesis degradation, Birth death stochastic process, Analytic calculations, Generating function

1. Introduction

Proteins are essential macromolecules that serve both as structural components of the cell and as its enzymatic machinery. When stochasticity in this processes is ignored the deterministic Michaelis-Menten model which can be understood also as a mean field approach is a good approximation. In most cases stochasticity plays a significant role in the process that can not be ignored. For example gene expression in both prokaryotes and eukaryotes is inherently stochastic [1, 2, 3, 4]. Also, due to the small number of molecules involved, the gene regulation by transcription factor proteins is a stochastic process. In general, by measuring the intensity profiles of fluorescence markers, we are able to directly observe fluctuations in the concentrations of proteins and mRNAs [5,6]. Such stochasticity is both controlled and exploited by the cells, therefore must be included in models.

Protein synthesis is a tightly gene expression regulated cellular process that affects growth, reproduction, aging and survival in response to both intrinsic and extrinsic cues, such as nutrient availability and energy levels. The turnover (synthesis) of these proteins is a stochastic process that plays a critical role in all biological processes.

Most of the modeling activity in regard with protein-protein interactions is using stochastic simulations and various numerical approaches that require certain approximations [7-11]. By developing analytical approaches, to the problem of protein interactions, a new understanding can be revealed.

The aim of this work is to explore minimal models of protein synthesis degradation and gain some analytical insight. For clarity of exposition, in the next section we present the analytic tools used to find the steady state probability distributions of protein copy numbers in

the cell and introduce the model in its general form. After that we are illustrating on an concrete case and find an analytical solution.

* Corresponding author: dianadavidrus@gmail.com; diana.david-rus@nipne.ro

2. Methods and general framework of the model

We envision a protein synthesis degradation process as a continuous in time birth death Markov process with a discrete (very large) state space. When two different species or types of proteins with a large number of possible states (or copies) are involved, the stochastic model that describes such process can be considered a two dimensional birth death Markov process.

For a simple case where each reaction either creates or annhilates one and only one component, and the birth/decay rates are constant, the master equation describing the time evolution of the probability distribution gives the flow for $\pi(j, k; t) = \pi_{jk}$, the probability of there being j copies of the 1st species, k copies of the second, at time t.

$$\frac{d\pi_{jk}}{dt} = \sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\left(\pi_{j-1k}\beta + \pi_{jk-1}b_{jk} - \pi_{jk}j\delta - \pi_{jk}kd\right) \quad (1)$$

π_{jk} is the joint probability distribution of type I and II to have j respective k copies; δ and d are the rates of degradation of protein type 1 respectively 2, some constants proportional to the existing number of protein copies. The rates of creation for protein type 1 with j copies is β and for protein type 2 with k copies is b_{jk} , (see Fig. 1).

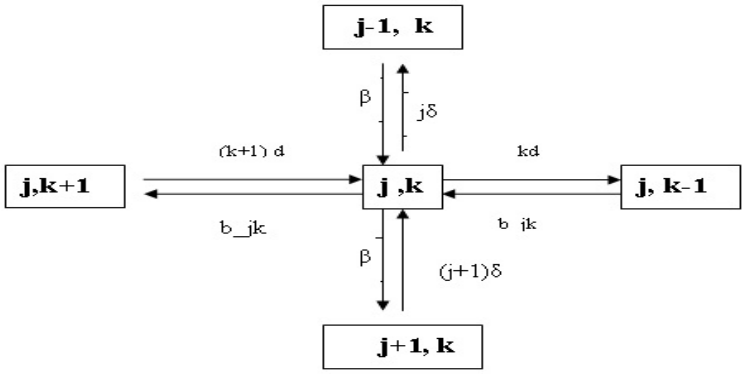

Fig.1 synthesis-degradation of protein type I,II with j, respectively k copies, β, bjk - the rate of synthesis for protein type I respectively II, (j and (j+1))δ, (k and (k+1))d - the rate of degradation for protein type I respectively II, proportional with the number of protein copies that are degradating.

The interaction between the two protein types, I and II is described by a step function as following: when protein type 1 gets to a certain threshold θ, protein type 2 is changing the rate of creation from b_0 to b_1

$$b_{jk} = \begin{cases} b_0, \text{when } j < \theta \\ b_1, \text{when } j \geq \theta \end{cases}$$

We will study analytically the stochastic model formulated above in the steady state case for a particular choice of states and rules of state transitions.

Solving master equation (1) analytically for the long time behavior of π_{jk} is generally an impossible task when the state space is very large. One, therefore, has to resort on various techniques. One such technique often used successfully in stochastic processes literature is the

"generating function technique" [12-14]. We remind the reader of some well-known aspects of this technique in order to make the present discussion self-contained.

Assume j is a discrete random variable and assume, for convenience, the state space is {0, 1, 2...}. Let π_j be the probability mass function of j where $\sum_{j=0}^{\infty}\pi_j$ =1; the mean of j satisfy: $\mu_j = E(j) = \sum_{j=0}^{\infty} j\pi_j$. The probability generating function (p.g.f.) of the discrete random variable j is defined by

$$f_j(x) = E(x^j) = \sum_{j=0}^{\infty} \pi_j x^j$$

for some $x \epsilon R$, because $\sum_{j=0}^{\infty}\pi_j$ =1, the above sum converges absolutely for $|x| \leq 1$. As the name implies, the p.g.f. generates the probabilities associated with the distribution, where $f_j(0) = \pi_0$, $f_j'(0) = \pi_1$,

$f_j''(0) = 2!\pi_2$ and in general $f_j^n(0) = n!p_n$ .The p.g.f. gives entire information associated with the distribution. Another known result that we will use and therefore remind here is that for the case when we have just one protein type undergoing a birth death process with a constant birth (β)/decay (δ) rate, is a well known fact that in steady state, its stationary probability distribution is a Poisson distribution, i.e, $p_j = \frac{1}{j!}\left(\frac{\beta}{\delta}\right)^j e^{-\beta/\delta}$, see [13,14].

3. Results: minimal model considered - a two dimensional birth-death process

Given that real biological systems frequently involve small numbers of molecules we developed a minimal model where the second protein type can be in 2 possible states: present or absent. An example of such situation in real life would be a genetic switch on/off.

The two protein types undergo a birth/death process with interaction. The creation and anihilation rates are the same as in general model with the following specifications: type I protein can have any number of copies/states, while type II protein can have only 2 possible copies/states: 0 or 1, meaning we have no protein or just one protein. The interaction between the 2 protein types is as following: the creation rates of the second protein type, b_{jk} , will directly depend by the number of copies of first protein type as bellow:

$$b_{jk} = \begin{cases} b_0, when\ j \leq \theta \\ b_1\ when\ j > \theta \\ 0\ when\ k > 1\ \forall j \end{cases}$$

Note that in this model for simplicity θ = 1, see fig. 2. The result can be easily generalized for any θ threshold.

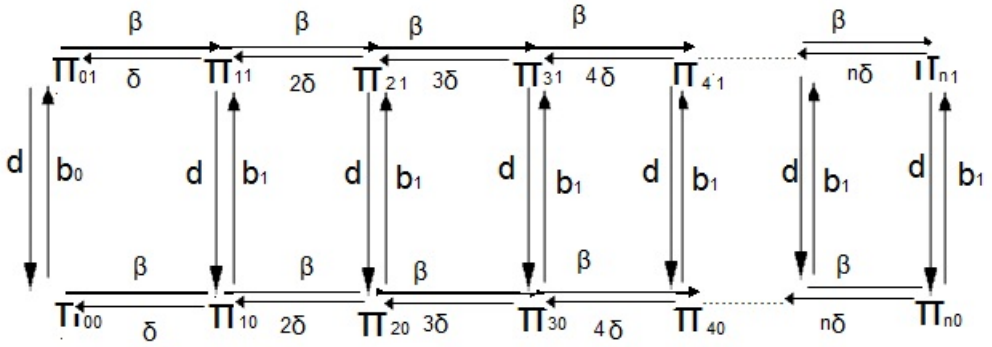

Fig. 2 Synthesis-Degradation of the two protein types as a two dimensional birth-death problem; Πjk is the joint probability distribution of protein type I and II, where j=1...n; k=0,1. The creation rate of protein type II is changing from b0 to b1 when number of protein copies type I is reaching the threshold θ (for simplification here θ = 1)

For this model, in steady state the master equation (1) is replaced with bellow equations, where we keep the same notation as in equation (1).

$$j = 0,\ k = 0 : \pi_{00}(\beta + b_0) = \pi_{10}\delta + \pi_{01}d, \tag{2}$$

$$j \geq 1,\ k = 0 : \pi_{j0}(j\delta + \beta + b_1) = \pi_{j+1;0}(j+1)\delta + \pi_{j1}d + \pi_{j-1;0}\beta, \tag{3}$$

$$j = 0,\ k = 1 : \pi_{01}(\beta + d) = \pi_{11}\delta + \pi_{00}b_0, \tag{4}$$

$$j \geq 1,\ \ k = 1 : \pi_{j1}(j\delta + \beta + d) = \pi_{j+1,1}(j+1)\delta + \pi_{j0}b_1 + \pi_{j-1,1}\beta, \tag{5}$$

Using generating function technique we simplify our problem. The above equations are transformed into ODE's satisfied by generating function.

Let $f_k(x) = \sum_{j=0}^{\infty} \pi_{jk}x^j$ be the probability generating function (p.g.f); the differential of p.g.f. is: $f_k'(x) = \sum_{j=0}^{\infty} j\pi_{jk}x^{j-1}$; where $\sum_{j=0}^{\infty} \pi_{jk}x^j$ converges absolutely for $|x| \leq 1$. According with the above notation, we have: $f_k(0) = \pi_{0k}$, $f_0(0) = \pi_{00}$, $f_1(0) = \pi_{01}$ and $f_0(x) = \sum_{j=0}^{\infty} \pi_{j0}x^j$, $f_1(x) = \sum_{j=0}^{\infty} \pi_{j1}x^j$.

Given that, for the case when we have just one protein type undergoing a birth death process with a constant birth/decay rate in steady state, it's stationary probability distribution is a Poisson distribution, and since protein 2 doesn't influence protein 1, the marginal distribution of protein 1 will still be given by the Poisson distribution: $p_j = \pi_{j0} + \pi_{j1} = \frac{1}{j!}\left(\frac{\beta}{\delta}\right)^j e^{-\beta/\delta}$. It follows that marginal, in the generating function notation is correct to right:

$$\mathrm{f_0(x) + f_1(x)} = \sum_{\mathrm{j=0}}^{\infty} \mathrm{x^j}\left(\pi_{\mathrm{j0}} + \pi_{\mathrm{j1}}\right) = \sum_{\mathrm{j=0}}^{\infty} \mathrm{x^j}\, e^{-\beta/\delta}\, \frac{1}{j!}\left(\frac{\beta}{\delta}\right)^j =$$

$$= e^{-\beta/\delta} \sum_{\mathrm{j=0}}^{\infty} \frac{1}{j!}\left(x\frac{\beta}{\delta}\right)^j = e^{-\beta/\delta}e^{-x\beta/\delta} = e^{(x-1)\beta/\delta} \tag{6}$$

And

$$f_0'(x) + f_1'(x) = \frac{\beta}{\delta}e^{(x-1)\beta/\delta} \tag{7}$$

Applying generating function technique on equations 2,3 an ODE equation (eq.8) satisfied by a generating function is derived (see Appendix A for details in derivation)

$$x\delta\frac{d}{dx}f_0(x) + (\beta + b_1)f_0(x) + (b_0 - b_1)f_0(0) =$$

$$= \delta\frac{d}{dx}f_0(x) + f_1(x)d + f_0(x)\beta x \tag{8}$$

same procedure applied on eq 4,5 and obtain the following equation:

$$x\delta\frac{d}{dx}f_1(x) + (\beta + d)f_1(x) =$$

$$= (b_0 - b_1)f_0(0) + \delta\frac{d}{dx}f_1(x) + f_0(x)b_1 + xf_1(x)\beta \qquad (9)$$

Steps toward obtaining $f_0(x)$: From condition (6) and eq.(8) we obtain a new equation (eq.10) in the $f_0(x)$ as unknown.

$$(x-1)\delta\frac{d}{dx}f_0(x) + [-\beta(x-1) + b_1 + d]f_0(x) =$$

$$= de^{(x-1)\beta/\delta}(b_1 - b_0)f_0(0) \qquad (10)$$

Solving this equation one will get the analytical expression for generating function $f_0(x)$. This is a first order ODE. Using integrand factor method one gets after some calculations, eq. 11(see appendix B for details in derivation).

$$f_0(x)(1-x)^{(b_1+d)/\delta}e^{-\beta/\delta x} - \mathrm{f}_0(0) =$$

$$= -de^{-\beta/\delta}\frac{1}{b_1+d}\left[1-(1-x)^{(b_1+d)/\delta}\right] - \frac{(b_1-b_0)f_0(0)}{\delta}\int_0^x e^{-\frac{\beta}{\delta}y}\,(1-y)^{\left(\frac{b_1+d}{\delta}-1\right)}dy \qquad (11)$$

$$-$$

Setting $x = 1$ in eq. (11) one obtains $f_0(0)$:

$$f_0(0) = \frac{de^{-\frac{\beta}{\delta}}}{b_1+d}\left[\frac{1}{1-\frac{(b_1-b_0)}{\delta}\int_0^1 e^{-\frac{\beta}{\delta}y}\,(1-y)^{\left(\frac{b_1+d}{\delta}-1\right)}dy}\right] \qquad (12)$$

Going back at condition (6): $\mathrm{f}_1(\mathrm{x}) = e^{(x-1)\beta/\delta} - f_0(x)$, and setting
$x = 0$ we get : $\mathrm{f}_1(0) = e^{-\beta/\delta} - f_0(0)$
Knowing that the derivatives of generating fct at zero gives the probabilities associated with the distribution, we have:

$\pi_{00} = f_0(0)$*;* $\pi_{10} = f_0'(0)$*;*$\pi_{20} = f_0''(0)$*;----------;*$\pi_{n0} = f_0^n(0)$
and $\pi_{01} = f_1(0)$*;*$\pi_{11} = f_1'(0)$*;*$\pi_{21} = f_1''(0)$*;* *--------;* $\pi_{n1} = f_1^n(0)$

and as such, one has access to entire distribution.

4. Discussions and conclusions

Since stochasticity plays a major role in biological processes , I am developing methods for solving/describing two dimensional stochastic processes that involve interactions. The model involves the "simple" but biologically important problem of protein interactions and stochastic interactions therein. I am building an artificial toy model that describes a two dimensional generation/degradation process of two different types of proteins interacting with each other in the following manner: the rate of generation of one type of protein changes once the quantity of the second protein is above a certain threshold. I use the method of dimensionality reduction for approaching this problem and getting an analytical expression for the joint probability distribution.

Using generating function technique I've shown how one can get an analytical expression for joint probability distribution of the two protein types that undergo a birth/death process with interaction. The analytical result can be used to help developing appropiate numerical methods for approximating results of the stochastic process when the second protein type has more states. Given that biochemical processes frequently involve small numbers of molecules (e.g. a few molecules of a transcriptional regulator binding to one 'molecule' of a DNA regulatory region) and such reactions are subject to significant stochastic fluctuations we consider that the minimal model considered here is relevant to such context and therefore important to be studied at analytical level.

Overall this relatively simple model can be used to evaluate the impact of stochastic factors in protein folding on biological fitness. Such analysis constitute a core unit for considering the complexities of multiple stochastic processes that are relevant for protein- protein interactions.

Acknoledgements

This work would have not been possible without the guidance of Prof. Larry Shepp and the valuable insights of Prof. Joel L. Lebowitz. Note that this result was previously communicated to Dr. Chris Wiggins who challenged us with the problem of identifying an analytical expression. This work was partly supported by NSF Grant DMR-0802120 and AFOSR Grant AF-FA-9550-04-4-22910. Also by POSDRU/89/1.5/S/60746.21 and CNCSIS-UEFISCDI, project number PN-II-PTPCCA-2011-3.1-1350.

References

[1] Maheshri, O'Shea, E.K., Annu. Rev. Biophys. Biomol. Struct. **36**,413 (2007).
[2] Elowitz MB, Levine AJ, Siggia ED, Swain PS, Science **297**,1183 (2002)
[3] Blake WJ, Kaern M, Cantor CR, Collins JJ, Nature **422**, 633 (2003)
[4] Raser JM, O'Shea EK, Science **304**,1811(2004)
[5] Kaern, M., Elston, T.C., Blake, W.J. Collins, J.J. Nat. Rev. Genet. **6**, 451 (2005).
[6] Kaufmann, B.B. van Oudenaarden, A, Curr. Opin. Genet. Dev. **17**, 107 (2007).
[7] J. S. van Zon and P. R. ten Wolde, Physical Review Letters **94(**128), 103 (2005).
[8] B. Munsky and M. Khammash, Journal of Computational Physics **226**, 818 (2007).
[9] S. Lampoudi, D. T. Gillespie, and L. R. Petzold, J. Chem. Phys. **130**, 104 (2009).
[10] D. Gillespie, M. Roh, and L. Petzold, J. Chem. Phys. **130**, 174 (2009).
[11] Brian Drawert, Michael J. Lawson, Linda Petzold, and Mustafa Khammash, J. Chem. Phys. **132**, 741(2010)
[12] Bailey N.T.J, The Elements of Stochastic Processes with Applications to the Natural Sciences, J. Wiley and Sons, New York, (1990)
[13] Karlin,S. and H.Taylor, First and Second Course in Stochastic Processes, (1975, 1981)
[14] Linda S. Allen,Stochastic processes with applications to Biology, (2003)

7. Appendix

A . Apply generating function technique in two state model for obtaining an ODE satis_ed by g.f.: eq. (8)

Using generating function technique on equations 2,3 (in equation 2+3 multiply by x^j and sum over j from 0 to ∞ each term) an ODE equation (eq.8) satisfied by a generating function is derived.

In equation 2+3 multiply by x^j and sum over j from 0 to ∞ each term.

$$\delta \sum_{j=1}^{\infty} j\pi_{j0}x^j + \sum_{j=1}^{\infty}(\beta + b_1)\pi_{j0}x^j + \beta\pi_{00} + b_0\pi_{00} =$$

$$= \delta \sum_{j=1}^{\infty}(j+1)\pi_{(j+1),0}x^j + d\sum_{j=1}^{\infty}\pi_{j1}x^j + \beta\sum_{j=1}^{\infty}\pi_{(j-1),0}x^j + \delta\sum_{j=1}^{\infty}\pi_{10}x^j + + d\sum_{j=1}^{\infty}\pi_{01}x^j$$

where each term in the equation can be written:

I

$$\delta \sum_{j=1}^{\infty} j\pi_{j0}x^j = x\delta\sum_{j=0}^{\infty} j\pi_{j0}x^{j-1} = x\frac{d}{dx}f_0(x)\delta$$

II

$$\sum_{j=1}^{\infty}(\beta + b_1)\pi_{j0}x^j - (\beta + b_1)\pi_{00}x^0 = (\beta + b_1)\mathrm{f}_0(\mathrm{x}) - (\beta + b_1)\mathrm{f}_0(0)$$

.

III

$$\beta\pi_{00} = \beta f_0(0)$$

IV

$$b_0\pi_{00} = b_0 f_0(0)$$

V

$$\delta \sum_{j=1}^{\infty}(j+1)\pi_{(j+1),0}x^j = \delta\sum_{j=0}^{\infty}(j+1)\pi_{(j+1),0}x^j - \delta\pi_{10}$$

VI

$$d\sum_{j=1}^{\infty}\pi_{j1}x^j = d\sum_{j=0}^{\infty}\pi_{j1}x^j - d\pi_{01} = f_1(x) - d\pi_{01}$$

VII

$$\beta\sum_{j=1}^{\infty}\pi_{(j-1),0}x^j = \beta x f_0(x)$$

VIII $\delta\pi_{10}$, IX $d\pi_{01}$

introducing all terms from I- IX back in equation we obtain equation (8):

$$x\delta\frac{d}{dx}f_0(x) + (\beta + b_1)f_0(x) + (b_0 - b_1)f_0(0) = \delta\frac{d}{dx}f_0(x) + f_1(x)d + f_0(x)\beta x$$

Similarly, applying generating function technique on eq 4,5 yields the following equation:

$$x\delta\frac{d}{dx}f_1(x) + (\beta + d)f_1(x) = (b_0 - b_1)f_0(0) + \delta\frac{d}{dx}f_1(x) + f_0(x)b_1 + xf_1(x)\beta \quad (9)$$

Further, using expression (6) in eq.(8) I obtain equation (eq.10) for $f_0(x)$

(10)

$$(x-1)\delta\frac{d}{dx}f_0(x)+[-\beta(x-1)+b_1+d]f_0(x)=de^{(x-1)\beta/\delta}(b_1-b_0)f_0(0)$$

B Derivations of eq.8 in two state model-Integrating factor method

$$(x-1)\delta\frac{d}{dx}f_0(x)+[-\beta(x-1)+b_1+d]f_0(x)=de^{(x-1)\beta/\delta}(b_1-b_0)f_0(0)$$

multiply above eq. with the integrand factor:

$$\left[(x-1)\delta f_0^{'}(x)+(-\beta(x-1)+b_1+d)f_0(x)\right]e^{\int_0^x\frac{b_1+d-\beta(y-1)}{(y-1)\delta}dy}=$$

$$(x-1)\delta\frac{d}{dx}\left(f_0(x)e^{\int_0^x\frac{b_1+d-\beta(y-1)}{(y-1)\delta}dy}\right)=$$

$$=\left(de^{(x-1)\beta/\delta}+(b_1-b_0)f_0(0)\right)e^{\int_0^x\frac{b_1+d-\beta(y-1)}{(y-1)\delta}dy}$$

where:

$$\int_0^x\frac{b_1+d-\beta(y-1)}{(y-1)\delta}dy=-\frac{\beta}{\delta}x+\ln(1-x)^{\frac{b_1+d}{\delta}}$$

therefore:

$$e^{\int_0^x\frac{b_1+d-\beta(y-1)}{(y-1)\delta}dy}=(1-x)^{\frac{b_1+d}{\delta}}\ e^{-\left(\frac{\beta}{\delta}\right)x}$$

from here we get by dividing with (x - 1)δ:

$$\left(\frac{d}{dx}f_0(x)e^{\int_0^x\frac{b_1+d-\beta(y-1)}{(y-1)\delta}dy}\right)=\frac{1}{(x-1)\delta[de^{(x-1)\beta/\delta}(b_1-b_0)f_0(0)]e^{-\left(\frac{\beta}{\delta}\right)x}e^{\left(\frac{b_1+d}{\delta}\right)}\log(1-x)}$$

Integrating this eq. we obtain:
(10)

$$f_0(x)e^{\int_0^x\frac{b_1+d-\beta(y-1)}{(y-1)\delta}dy}-f_0(0)=$$

$$=-\int_0^x\left(de^{(y-1)\beta/\delta}+(b_1-b_0)f_0(0)\right)e^{-\left(\frac{\beta}{\delta}\right)y}\frac{1}{\beta}\frac{1}{(1-y)^{\left[1-\left(\frac{b_1+d}{\delta}\right)\right]}}$$

or equation. (11.a):

$$f_0(x)(1-x)^{\frac{b_1+d}{\delta}}\ e^{-\left(\frac{\beta}{\delta}\right)x}-f_0(0)=I_1+I_2$$

where

$$I_1 = -\frac{d}{\delta}\int_0^x e^{\frac{\beta}{\delta(y-1)}} e^{-\left(\frac{\beta}{\delta}\right)y} \frac{1}{(y-1)^{\frac{-b_1+d}{\delta}+1}} dy =$$

$$= -de^{-\left(\frac{\beta}{\delta}\right)} \frac{1}{b_1+d}\left[1-(1-x)^{(b_1+d)/\delta}\right]$$

and

$$I_2 = -\frac{(b_1-b_0)f_0(0)}{\delta}\int_0^x e^{-\left(\frac{\beta}{\delta}\right)y}(1-y)^{\left(\frac{b_1+d}{\delta}-1\right)} dy$$

by introducing the expression I_1 and I_2 in equation [11] we obtain the expression for $f_0(x)$:

[11] $$f_0(x)(1-x)^{\frac{b_1+d}{\delta}}\ e^{-\left(\frac{\beta}{\delta}\right)x} - f_0(0) =$$

$$= -de^{-\left(\frac{\beta}{\delta}\right)} \frac{1}{b_1+d}\left[1-(1-x)^{(b_1+d)/\delta}\right] -$$

$$-\frac{(b_1-b_0)f_0(0)}{\delta}\int_0^x e^{-\left(\frac{\beta}{\delta}\right)y}(1-y)^{\left(\frac{b_1+d}{\delta}-1\right)} dy$$