Some New Results in the Mathematical Theory of Phage-Reproduction

by

Prem S. Puri*

Purdue University, Lafayette, Indiana

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 146

March 1968

*This investigation was supported (in part) by the U.S. Public Health Service, National Institute of Health Grant GM-10525, at the Statistical Laboratory, University of California, Berkeley.

Some New Results in the Mathematical Theory of Phage-Reproduction

by

Prem S. Puri

Purdue University, Lafayette, Indiana

0. Summary. In the theory of phage reproduction, the mathematical models considered thus far (see Gani [5])assume that the bacterial burst occurs a fixed time after infection, or after a fixed number of generations of phage-multiplications, or when the number of mature bacteriophages reaches a fixed threshold. In the present paper, such hypotheses of fixed thresholds are abondoned in favour of a more realistic assumption: Given that until the time t the bacterial burst has not yet taken place, the occurrence of the burst between t and $t + \Delta t$ is treated as a random event, the probability of which is $f(\cdot|t)\Delta t + o(\Delta t)$, where f is a non-negative and nondecreasing function of the number $X(t)$ of vegetative phages and of $Z(t)$, the number of mature bacteriophages at time t. More specifically it is assumed that $f = b(t)\, X(t) + c(t)\, Z(t)$ with $b(t), c(t) \geq 0$. Here $X(t)$ is assumed to be a linear birth and death process and $Z(t)$ corresponds to the number of deaths until time t. One of the problems considered here is the joint distribution of X_T and Z_T, the numbers at burst of vegetative and mature bacteriophages respectively. The distribution of Z_T is then fitted to observed data due to Delbrück [2].

1. Introduction. The present work has emerged as a result of inspiration and stimulation the author received first from an interesting paper of Professor Gani [5], where he gives an excellent account of various stochastic models for

*This investigation was supported (in part) by the U.S. Public Health Service, National Institute of Health Grant GM-10525, at the Statistical Laboratory, University of California, Berkeley.

bacteriophage put forth thusfar; second from Professor Hirsch of Courant Institute of Mathematical Sciences who pointed out during the author's visit to the institute about the possibility of using author's methods of [14] in developing an improved model of phage-reproduction. The author is indebted for this to both Professor Gani and Professor Hirsch.

The aim of this paper is to develop a stochastic model of the fairly well known mechanism concerning the phage-reproduction and its parasitic cycle of growth while feeding on the bacterium (see Gani [5]). A brief but simplified sketch of the mechanism is desirable here. After the initial insertion of the DNA strand (vegetative phage) into the bacterium, a period of length τ (possibly random but varying from seven to ten minutes), known as eclipse, follows. During this period, the vegetative phage produces within the bacterium a random number of its copies according to a birth process, and simultaneously the parts necessary for the conversion of a vegetative phage to a mature phage are under production. Furthermore, the infected bacterium is no longer capable of reproduction and is considered as dying immediately after infection until the lysis (burst) that occurs following eclipse. During the period between τ and $(T + \tau)$, the time of occurrence of the lysis, while some vegetative phages produce more of their copies (births) there are others which are rendered inactive after they emerge as mature phages. This latter process between τ and $(T + \tau)$ has been studied as a birth and death process $X(t)$. (See Steinberg and Stahl [15] and Gani [4]). Here the death corresponds to the conversion of a vegetative phage to a mature one, and $X(t)$ stands for the number of vegetative phages at time t.

One of the problems of interest is the study of distributions of the burst time $(T + \tau)$ and the burst size Z_T, that is, the number of <u>mature</u> phages released at $(T + \tau)$, the time of the lysis. These distributions have been studied by

several research workers under different assumptions. Steinberg and Stahl [15] and Gani [4] have investigated the distribution of Z_T under the assumption that the burst time $(T + \tau)$ is fixed. Others (see Gani [5] for references) have studied the distribution of burst time assuming that the burst takes place as soon as the number of mature phages $Z(t)$ reaches a fixed threshold. Kimball [8] has considered a discrete time model where he assumes that the phage reproduction takes place in generations, that for k generations after infection the phage remains vegetative, and thereafter mutation occurs at each subsequent generation with probability half until lysis at the $(k + m)$th generation; both k and m are assumed constant. The problem of interest here is the burst size distribution. Kimball's model although explains some aspects of the observed distribution but fails to explain others. Unfortunately, most of the models proposed above have never been tried on observed data. This may be attributed partly to the lack of reliable data and partly to the algebraically messy form of the theoretical distributions obtained under various models.

We note that the hypothesis of existence of a fixed threshold of the number of mature phages (burst size) or of a fixed burst time is a common feature of all the models considered thus far. In the present paper this hypothesis is abandoned, for it is well known that neither T nor Z_T is fixed and that in reality they appear to be random variables with some probability distribution. Instead, we adopt here a more realistic assumption originally suggested by Professor LeCam and used elsewhere by the author [14]. Given that the bacterial burst has not taken place during $(0,t]$, the occurrence of the burst between t and $t + \Delta t$ is treated as a random event, the probability of which can be written as $f(\cdot|t)\Delta t + o(\Delta t)$, where f is a nonnegative and nondecreasing function of $X(t)$ and $Z(t)$. With this treatment, both T and Z_T become random variables; the deduction of their

joint distribution is one of the problems treated in this paper. We shall also see how the present model under certain conditions explains the observed phenomenon (see Delbrück [2]) of similarity of the burst size distribution in the case of single infection to the one for the multiple infection case.

2. A Stochastic Model for Phage Reproduction and the Bacterial Burst.

We assume the origin on the time scale to be the moment when the bacterium is first infected. Let $X(t)$ be the number of vegetative phages within the bacterium at time t, with $X(0)$ equal to m, the number of phages that infected the bacterium at time zero. Although, in principle, the bacterium may get reinfected any number of times during the eclipse period $(0,\tau)$ after the initial infection, it is still reasonable to assume that no further infection takes place after the initial one at time zero. We also assume that the length τ of the eclipse period is fixed. Let $(T + \tau)$ be the moment of time when the lysis takes place. Following Gani [5] the process $X(t)$ is considered here to be a linear birth and death process with birth and death rates $\tilde{\lambda}(t)$ and $\tilde{\mu}(t)$ given by

$$\tilde{\lambda}(t) = \begin{cases} \nu(t) & 0 < t \leq \tau \\ \lambda(t) & t > \tau \end{cases} \tag{1}$$

and

$$\tilde{\mu}(t) = \begin{cases} 0 & 0 < t \leq \tau \\ \mu(t) & t > \tau \end{cases}, \tag{2}$$

respectively. Furthermore, let $f(X(t), Z(t), t)$ be the risk function for the lysis of the bacterium so that

$$\Pr[\text{lysis takes place during } (t,t+\Delta t) \mid X(t)=x,\ Z(t)=z,\ Y(t)=1] = [f(x,z,t)]\,\Delta t + o(\Delta t)\ , \tag{3}$$

where $Z(t)$ is the number of mature phages (deaths) at time t, and

$$Y(t) = \begin{cases} 1 & \text{if lysis has not occurred until time } t \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

As a first attempt, we assume for the sake of simplicity, that

$$f(x,z,t) = \begin{cases} 0 & 0 < t \le \tau \\ b(t)x + c(t)z\ , & t > \tau\ , \end{cases} \tag{5}$$

where $b(t)$ and $c(t)$ are arbitrary nonnegative bounded functions. Again, since $\mu(t) = f(x,z,t) = 0$ for $0 < t \le \tau$, we have $Y(\tau) = 1$ and $Z(\tau) = 0$. Thus we can conveniently consider to begin with, the process $\{X(t), Z(t), Y(t)\}$ conditionally for $t > \tau$, given $X(\tau)$, $Z(\tau) = 0$ and $Y(\tau) = 1$. Given $X(\tau) = n$, let

$$G_1(u,v;\ t|n) = \sum_{x=0}^{\infty}\sum_{z=0}^{\infty} u^x v^z\, p_{x,z,1}(t|n);\ |u|,\ |v| \le 1\ , \tag{6}$$

for $t > \tau$, where for $x,z = 0,1,2,\ldots,$

$p_{x,z,1}(t|n) = \Pr[X(t) = x,\ Z(t) = z,\ Y(t) = 1 \mid X(\tau) = n,\ Z(\tau) = 0,\ Y(\tau) = 1]$.

Following Puri [14], it can be shown that the probability generating function (p.g.f.) $G_1(u,v,t)$ satisfies the partial differential equation

$$G_t = [\lambda(t)u^2 - (\lambda(t)+\mu(t)+b(t))u+\mu(t)v]\,G_u - c(t)\,v\,G_v\ , \tag{7}$$

subject to the initial condition

$$G_1(u,v;\tau) = u^n \ , \tag{8}$$

where G_t, G_u and G_v denote the corresponding first order partial derivatives of G_1. Equation (7) appears difficult to solve in this generality. However, it can be solved for the case analogous to the one considered by Gani and Yeo [6], where we assume that for all $t > \tau$,

$$\frac{\mu(t)}{\lambda(t)} = \rho \ , \ \frac{b(t)}{\lambda(t)} = \delta \ , \ \frac{c(t)}{\lambda(t)} = \theta \ , \tag{9}$$

with ρ, δ and θ as some nonnegative constants. Subject to (9), the equation (7) becomes

$$[\lambda(t)]^{-1} G_t = [u^2 - (1+\rho+\delta)\, u + \rho v]\, G_u + \theta v G_v = 0 \ . \tag{10}$$

The auxiliary equations associated with (10) are given by

$$\lambda(t)dt = \frac{du}{(1+\rho+\delta)u - u^2 - \rho v} = \frac{dv}{\theta v} = \frac{dG_1}{0} \ , \tag{11}$$

from which we have,

$$\frac{dv}{dt} = \lambda(t)\ \theta\ v$$

whence

$$v = C_1 \exp\{\theta \int_\tau^t \lambda(s)ds\} \ ; \ C_1 = v \exp\{-\theta \int_\tau^t \lambda(s)ds\} \ , \tag{12}$$

where C_1 is the constant of integration. Substituting for v from (12) in the auxiliary equations, we have

$$\frac{du}{dt^*} = (1 + \rho + \delta)\, u - u^2 - \rho\, C_1 \exp[\theta t^*] , \tag{13}$$

where $t^* = \int_\tau^t \lambda(s)\, ds$. The problem here is to solve equation (13) which is similar to the one arose elsewhere (see Puri [13]) . Substituting $\xi = u - \frac{1}{2}(1+\rho+\delta)$ in (13), we have

$$\begin{aligned} \frac{d\xi}{dt^*} &= -\xi^2 + \left(\frac{1+\rho+\delta}{2}\right)^2 - \rho\, C_1 \exp[\theta t^*] \\ &= \frac{h'(t^*)}{g(t^*)}\, \xi^2 - \frac{g'(t^*)}{h(t^*)} , \end{aligned} \tag{14}$$

say, where the functions h and g satisfy the relations

$$\begin{cases} \dfrac{h'(t^*)}{g(t^*)} = -1 \\ \dfrac{g'(t^*)}{h(t^*)} = \rho\, C_1 \exp[\theta t^*] - \left(\dfrac{1+\rho+\delta}{2}\right)^2 . \end{cases} \tag{15}$$

Once (15) is solved for h and g , (14) immediately yields

$$\xi(t^*) = -\frac{g(t^*)}{h(t^*)} - \frac{1}{h^2(t^*)} \left(C_2 - \int_0^{t^*} \frac{h'(s)}{h^2(s)g(s)}\, ds\right)^{-1} , \tag{16}$$

where C_2 is the constant of integration. Eliminating g from (15) we have

$$h''(t^*) + \left[\rho\, C_1 \exp(\theta t^*) - \left(\frac{1+\rho+\delta}{2}\right)^2\right] h(t^*) = 0 . \tag{17}$$

With change of variable from t^* to w where

$$w = \frac{2}{\theta} (\rho\ C_1)^{\frac{1}{2}} \exp[\frac{\theta t^*}{2}] \quad , \tag{18}$$

(17) yields

$$w^2 \frac{d^2h}{dw^2} + w \frac{dh}{dw} + (w^2 - p^2)\ h = 0 \quad , \tag{19}$$

where $p = (1+\rho+\delta)/\theta$. Equation (19) is the well known Bessel equation and its solution is given by

$$h = J_p(w) = \sum_{k=0}^{\infty} \frac{(-1)^k\ (w/2)^{2k+p}}{k!\ \Gamma(k+p+1)} \quad . \tag{20}$$

Thus we have

$$h(t^*) = J_p(\frac{2}{\theta}\ [\rho\ C_1\ \exp(\theta t^*)]^{\frac{1}{2}}) \tag{21}$$

and

$$g(t^*) = [\rho\ C_1 \exp(\theta t^*)]^{\frac{1}{2}}\ J_p'(\frac{2}{\theta}\ [\rho\ C_1 \exp(\theta t^*)]^{\frac{1}{2}}) \ . \tag{22}$$

Finally substituting these in (16) and using the fact that $u = \xi + \frac{1}{2}(1+\rho+\delta)$, we obtain solution of (13) as

$$u = \frac{1+\rho+\delta}{2} + \frac{J_p'\ (\frac{2}{\theta}\ [\rho\ C_1 \exp(\theta t^*)]^{\frac{1}{2}})\ [\rho\ C_1 \exp(\theta t^*)]^{\frac{1}{2}}}{J_p(\frac{2}{\theta}\ [\rho\ C_1 \exp(\theta t^*)]^{\frac{1}{2}})}$$

$$+ J_p^{-2}\ (\frac{2}{\theta}\ [\rho\ C_1 \exp(\theta t^*)]^{\frac{1}{2}})\ [C_2 + \int_0^{t^*} J_p^{-2}\ (\frac{2}{\theta}\ [\rho\ C_1 \exp(\theta s)]^{\frac{1}{2}}) ds]^{-1} . \tag{23}$$

Solving (23) for the constant C_2 we have

$$C_2 = -\int_0^{t^*} J_p^{-2}\left(\frac{2}{\theta}\left[\rho\, C_1 \exp(\theta s)\right]^{\frac{1}{2}}\right)ds + \left\{\left(u - \frac{1+\rho+\delta}{2}\right) J_p^2\left(\frac{2}{\theta}\left[\rho\, C_1 \exp(\theta t^*)\right]^{\frac{1}{2}}\right)\right. \tag{24}$$

$$\left. - \left[\rho\, C_1 \exp(\theta t^*)\right]^{\frac{1}{2}} J_p\left(\frac{2}{\theta}\left[\rho\, C_1 \exp(\theta t^*)\right]^{\frac{1}{2}}\right) J_p'\left(\frac{2}{\theta}\left[\rho\, C_1 \exp(\theta t^*)\right]^{\frac{1}{2}}\right)\right\}^{-1} .$$

Again from the auxiliary equations (11), we have, for $t > \tau$,

$$G_1(u,v;t|n) = \text{constant} = H(C_2) \quad , \tag{25}$$

say, where H is an arbitrary function. Using the initial condition (8) for determining the function H, we obtain after some manipulation

$$G_1(u,v;t|n) = \left[\frac{1+\rho+\delta}{2} + C_2^{-1} J_p^{-2}\left(\frac{2}{\theta}\left[\rho\, C_1\right]^{\frac{1}{2}}\right) + \left[\rho\, C_1\right]^{\frac{1}{2}} J_p'\left(\frac{2}{\theta}\left[\rho\, C_1\right]\right) J_p^{-1}\left(\frac{2}{\theta}\left[\rho\, C_1\right]\right)\right]^n ,$$

or, on substituting for C_2 and then for C_1 from (12) ,

$$G_1(u,v;t|n) = \left[\tilde{g}(u,v;\tau,t)\right]^n \quad , \tag{26}$$

where

$$\tilde{g}(u,v;\tau,t) = \left(\frac{1+\rho+\delta}{2}\right) + \frac{\theta x}{2} e^{-\frac{\theta t^*}{2}} \frac{J_p'\left(x\, e^{-\frac{\theta t^*}{2}}\right)}{J_p\left(x\, e^{-\frac{\theta t^*}{2}}\right)}$$

$$+ J_p^{-2}\left(x\, e^{-\frac{\theta t^*}{2}}\right)\left[\left\{\left(u - \frac{1+\rho+\delta}{2}\right) J_p^2(x) - \frac{\theta x}{2} J_p'(x)\, J_p(x)\right\}^{-1}\right.$$

$$\left. - \int_0^{t^*} J_p^{-2}\left(x\, e^{-\frac{\theta s}{2}}\right) ds\right]^{-1} \quad , \tag{27}$$

and

$$t^* = \int_\tau^t \lambda(s)\, ds \quad ; \quad x = \frac{2}{\theta}\sqrt{\rho v} \quad . \tag{28}$$

Again since $X(t)$ for $0 < t \leq \tau$ is a simple birth process, we have (see Kendall [7])

$$g(w;\tau) \equiv E(w^{X(\tau)}) = [h(w;\tau)]^m \quad , \tag{29}$$

where

$$h(w;\tau) = w \exp(-\int_0^\tau \nu(s)ds) \, [1 - w \{1 - \exp(-\int_0^\tau \nu(s)ds)\}]^{-1}. \tag{30}$$

Thus on replacing n in (26) by $X(\tau)$ and taking expectation over $X(\tau)$, we finally have for $t > \tau$

$$G_1(u,v;\tau,t) = g(\tilde{g}(u,v;\tau,t) = [h(\tilde{g}(u,v;\tau,t);\tau)]^m \quad , \tag{31}$$

where for $t > \tau$

$$G_1(u,v;\tau,t) \equiv \sum_{x=0}^{\infty} \sum_{z=0}^{\infty} u^x v^z \Pr[X(t) = x, Z(t) = z, Y(t) = 1 | X(0) = m, Z(0) = 0, Y(0) = 1]. \tag{32}$$

Remark 1. In the above calculations, it was assumed that $X(0) = m$ and that no further infection takes place after the initial one at time zero. Instead, if we let the origin of the time scale to be the moment when the bacterial suspension is added to that of bacteriophages so that $X(0) = 0$, and if for the eclipse period $0 < t < \tau$ it is assumed that

Pr[the bacterium gets infected by a single bacteriophage during $(t,t+h)$]$=L(t)h+o(h)$,

and

Pr[the bacterium gets infected by two or more bacteriophages during $(t,t+\tau)$]$=o(h)$,

where $L(t)$ is a given nonnegative bounded function defined on $0 < t < \tau$, then it is easy to establish that

$$g(w;\tau) = \exp[-\int_0^\tau L(s)\ \{1 - h(w;s,\tau)\}\ ds]\ , \tag{33}$$

where

$$h(w;s,\tau) = w\ \exp(-\int_s^\tau \nu(t)dt)[1 - w\{1 - \exp(-\int_s^\tau \nu(t)dt)\}]\ . \tag{34}$$

Using (33) in (31) one can easily obtain the expression for G_1 of (32) for this case.

Remark 2. While taking account of the experimental evidence for linear growth of the DNA content both during eclipse, and after, Ohlsen [9] has indicated the form of the function $\tilde{\lambda}(t)$ of (1) to be

$$\tilde{\lambda}(t) = \begin{cases} \dfrac{a}{(a-d)t+d\tau+1} & \text{for } t > \tau\ ; \\ \dfrac{a}{at+1} & \text{for } 0 < t \leq \tau\ . \end{cases} \tag{35}$$

Furthermore, based on this evidence, he observes that the ratio $\rho = \mu(t)/\lambda(t)$ of (9) should be constant for all $t > \tau$, and should approximately be equal to one. That $\rho \approx 1$ is based on the observation that on the average the number of vegetative phages does not change significantly after eclipse. (see Gani [5], pp. 232-33). Also, it does not appear inappropriate to assume that ρ, although is close to one, is strictly less than one. The assumption that the other ratios of (9) are also constant for $t > \tau$ is made only to enable us to solve equation (7) more easily.

Remark 3. Starting with $X(\tau) = n$, for a given realisation $\underset{\sim}{\omega}$ of $X(s)$ for all s with $\tau \leq s \leq t$, it is easy to establish that the probability of no burst during $(\tau,t]$ is given by

$$\exp\{-\int_{\tau}^{t} f(X(s,\underset{\sim}{\omega}),\ Z(s,\underset{\sim}{\omega}),\ s)ds\} . \tag{36}$$

Multiplying this by $u^{X(s,\underset{\sim}{\omega})}\ v^{Z(s,\underset{\sim}{\omega})}$ and taking expectation of the product, over all the realisations $\{X(s,\underset{\sim}{\omega});\ \tau \leq s \leq t\}$, we have G_1 of (6) as

$$G_1(u,v;t|n) = E[u^{X(t)}\ v^{Z(t)}\ \exp\{-\int_{\tau}^{t} f(X(s),Z(s),s)ds\}] . \tag{37}$$

For $f(x,z,s)$ satisfying (5), with $b(s)$ and $c(s)$ satisfying (9), (37) can be rewritten as

$$G_1(u,v;t|n) = E[u^{X(t)}\ v^{Z(t)}\ \exp\{-\delta\int_{\tau}^{t}\lambda(s)X(s)ds - \theta\int_{\tau}^{t}\lambda(s)Z(s)ds\}] . \tag{38}$$

Thus, as an interesting biproduct of the above theory, one observes, by simply treating δ and θ in (38) and (26) as nonnegative dummy variables, that the expression (26) is nothing but the transform (38) for the conditional joint distribution of the process $\{X(t),\ Z(t),\ \int_{\tau}^{t}\lambda(s)X(s)ds,\ \int_{\tau}^{t}\lambda(s)Z(s)ds\}$ for $t > \tau$, given that $X(\tau) = n$ and $Z(\tau) = 0$. The distribution problems concerning such random variables as integrals of certain stochastic processes have been studied elsewhere by Bartlett [1] and by the author (see [10,11,12]).

Finally the distribution of the time T, that the bacterium takes after the eclipse until its burst, is given by

$$1 - F_T(t) = Pr(T > t) = Pr(Y(t+\tau) = 1) = G_1(1,1;\tau,t+\tau) . \tag{39}$$

Furthermore, if the function $\lambda(t)$ is such that $\int_{\tau}^{\infty} \lambda(s)ds = \infty$, then we note from (27) and (31) that if $\theta > 0$,

$$\lim_{t\to\infty} G_1(1,1;\tau,t) = 0 , \tag{40}$$

so that T is an honest random variable. The case with $\theta = 0$, will be dealt separately in some detail in the next section.

3. Burst Size Distribution. In this section we shall attempt to find the joint distribution of the numbers of vegetative phages and mature bacteriphages at the time of burst of the bacterium. For $t_1 > t > \tau$, consider the p.g.f.

$$\phi(u,v,u_1,v_1;t,t_1)$$

$$= \sum_{z_1}\sum_{x_1}\sum_{z}\sum_{x} u^x v^z u_1^{x_1} v_1^{z_1} \Pr[X(t) = x,\ Z(t) = z,\ X(t_1) = x_1,$$

$$Z(t_1) = z_1,\ Y(t_1) = 1 | X(0) = m,\ Z(0) = 0,\ Y(0) = 1] , \tag{41}$$

where $|u|, |v|, |u_1|$, and $|v_1|$ are all less than or equal to one. Using (31), the expression for (41) turns out to be

$$\phi(u,v,u_1,v_1;t,t_1) = G_1(u\,\tilde{g}(u_1,v_1;t,t_1), v\,v_1;\ \tau,\ t) , \tag{42}$$

where $\tilde{g}(u_1,v_1;t,t_1)$ is given by (27) with (u,v,τ,t) replaced by (u_1,v_1,t,t_1).

Again, let X_T and Z_T denote respectively the numbers of vegetative and mature bacteriophages at the burst of the bacterium without regard to when it occurs. Also let

$$H_T(u,v) = \sum_{x=0}^{\infty} \sum_{z=0}^{\infty} u^x v^z \Pr(X_T = x, Z_T = z) . \tag{43}$$

Then following Puri [14], it can be easily shown that

$$H_T(u,v) = - \int_\tau^\infty \frac{d}{dt_1} \phi(u,v,1,1;t,t_1)\Big|_{t_1=t} dt . \tag{44}$$

Furthermore, the joint distribution of X_T, Z_T and T is given by

$$\Pr(X_T = x, Z_T = z, T = t) = \text{coefficient of } u^x v^z \text{ in} [-\frac{d}{dt_1} \phi(u,v,1,1;t,t_1)\Big|_{t_1=t}], \tag{45}$$

for $x,z = 0,1,2,\ldots$.

Unfortunately, in view of (27) and (31), the expressions for (44) and (45) become much too complicated. As such, we introduce at this point a simplifying assumption, namely that $C(t) \equiv 0$ or equivalently that $\theta = 0$. This implies that the risk of bacterial burst depends only on the number of vegetative phages and not on the mature ones. Although, in reality, this may not be strictly true, we shall see later that the simplified model based on this assumption does retain some of the elements necessary for its compatability with the observations. Furthermore, this may not appear unreasonable when one takes into account the fact that the cell material (DNA etc.) is getting used up by vegetative phages which multiply and not by the mature ones which do not. As such, the vegetative phages appear more important than the mature ones in weakening the bacterium and thereby leading eventually to a burst. Thus from hereon, it is assumed that $\theta = 0$. Also we assume that $\int_\tau^\infty \lambda(s)ds = \infty$.

For the case with $\theta = 0$, instead of simplifying the expression (27) , it is easier to solve (7) directly for $\tilde{g}(u,v;\tau,t)$ yielding

$$\tilde{g}(u,v;\tau,t) = G_1(u,v;t|X(\tau) = 1,\ Z(\tau) = 0,\ Y(\tau) = 1)$$

$$= \frac{r_1(u-r_2)\exp[-(r_1-r_2)\int_\tau^t \lambda(s)ds] - r_2(u-r_1)}{(u-r_2)\exp[-(r_1-r_2)\int_\tau^t \lambda(s)ds] - (u-r_1)} \quad , \tag{46}$$

where $r_1 = r_1(v)$ and $r_2 = r_2(v)$ are, with positive and negative signs, respectively,

$$\tfrac{1}{2}[(1 + \rho + \delta) \pm \{(1 + \rho + \delta)^2 - 4\rho v\}^{\frac{1}{2}}] \quad . \tag{47}$$

From (46) and (31) , one can easily obtain $G_1(u,v;\tau,t)$. Unfortunately with $\theta = 0$, T is not an honest random variable, since

$$\lim_{t\to\infty} \tilde{g}(u,v;\tau,t) = r_2(v) \quad ,$$

so that

$$\lim_{t\to\infty} P(T > t) = \lim_{t\to\infty} G_1(1,1;\tau,t+\tau) = [h(r_2^*\ ,\ \tau)]^m$$

$$= \left[\frac{r_2^*\ e^{-\int_0^\tau \nu(s)ds}}{1 - r_2^*\{1 - \exp(-\int_0^\tau \nu(s)ds)\}}\right]^m \quad , \tag{48}$$

where $r_2^* = r_2(1)$. Thus there is a positive probability, howsoever small it may be, that the bacterium never has a lysis after infection. This may very well be true in reality, even though the author is not aware of any experimental

evidence to this effect, except for a remark made by Delbrück ([2], see footnote on page 131) , where he makes the following statement. "We use the word 'infection' of a bacterium by a virus to designate the fact that a bacterium has adsorbed a virus particle. We do not imply that the infecting particle necessarily grows...." From this remark, one gets the impression that it is not improbable for the virus not to grow after infection and thereby leave the bacterium without a lysis. On the other hand, since the distribution of the burst size is meaningful only for those bacteria which do have bursts, the appropriate distribution for us to study is

$$\Pr(X_T = x,\ Z_T = z \,|\, T < \infty)\ ;\ x,z = 0,1,2,\ldots\ .$$

Let

$$H(u,v|m) = \sum_{x=0}^{\infty} \sum_{z=0}^{\infty} u^x v^z \Pr(X^T = x,\ Z_T = z \,|\, T < \infty);\ |u|,|v| \le 1\ , \tag{49}$$

be the corresponding p.g.f.. Then following Puri [14], we have

$$H(u,v|m) = \frac{-\int_{\tau}^{\infty} \frac{d}{dt_1} \phi(u,v,1,1;t,t_1)\Big|_{t_1=t} dt}{1 - [h(r_2^*\ ,\ \tau)]^m}\ . \tag{50}$$

The expression for ϕ , although easily obtainable using (42) and (46), is somewhat lengthy and will not be reproduced here. Substituting the expression for ϕ in (50) we obtain after some algebraic manipulation

$$H(u,v|m) = \left(\frac{u(1-r_1^*)(1-r_2^*)}{(u-r_1)(u-r_2)} \right) \left[\frac{\{h(u;\tau)\}^m - \{h(r_2;\tau)\}^m}{1 - \{h(r_2;\tau)\}^m} \right]\ . \tag{51}$$

By expanding (51) into a power series in u and v, one finds the individual probabilities $Pr(X_T = x,\ Z_T = z \mid T < \infty)$. Unfortunately, this process appears cumbersome; however, one can easily obtain moments of X_T and Z_T from (51), for instance

$$EX_T = \frac{1-\rho}{\delta} + \frac{m \exp(\int_0^\tau \nu(s)ds)}{1 - [h(r_2^*;\tau)]^m} \tag{52}$$

and

$$EZ_T = \frac{\rho}{\delta} - \frac{\rho}{(r_1^*-r_2^*)} \cdot \frac{m[h(r_2^*;\tau)]^{m-1}}{1-[h(r_2^*;\tau)]^m} \cdot \frac{\exp(-\int_0^\tau \nu(s)ds)}{[1-r_2^*\{1-\exp(-\int_0^\tau \nu(s)ds)\}]^2} \cdot \tag{53}$$

In order to see how far the joint distribution of X_T and Z_T given by (51) is compatible with the observations, one would need for various bacterial bursts, the observed number of both vegetative as well as the mature bacteriophages. Unfortunately, the author was unable to obtain any such data, possibly because the methods employed for bacteriophage counts (see Delbrück [2]) do not yield the exact burst time or the number of vegetative bacteriophages at the burst. Experimental counts for the burst size of mature bacteriophages are however possible and are available. In the next section, an attempt is made to fit the theoretical distribution of Z_T to the data published in the form of a histogram by Delbrück [2]. This is done after approximating the distribution of Z_T for small δ, under the seemingly appropriate assumption that $\rho < 1$.

4. The Fitting of the Theoretical Distribution of Z_T to Delbrück's Data.

We assume hereafter that ρ, although is close to one, is strictly less than one.

One of the features of the distribution of (X_T, Z_T), exhibited by (51), is that whereas $Pr(X_T = 0)$ is zero as expected, $Pr(Z_T = 0)$ is strictly positive and is given by

$$Pr(Z_T = 0) = \frac{\delta}{\rho+\delta} \frac{1}{1-[h(r_2^*;\tau)]^m} \quad . \tag{54}$$

This means that under the model with $\theta = 0$, there is a positive probability for a bacterial burst to yield no mature bacteriophage. This event appears highly unlikely, even though there is available no definite experimental evidence to this effect; the latter possibly being due to the fact that the methods of phage counts are such that they pick up only those bursts which have at least some mature bacteriophages present (see Delbrück [2]). How efficient these methods are particularly for bursts with very few mature phages is not clear to the author. Granting then, that the bursts with an absence of mature bacteriophages is a rare phenomenon, we conclude from (54) that among all the parameters, δ should be quite small. Also, since the distribution of (X_T, Z_T) as given by (51) appears rather involved, it may be relevant to approximate this for small values of the parameter δ. To this end, our next step is to find the limiting distribution of $(\frac{\delta}{1-\rho} X_T, \frac{\delta}{\rho} Z_T)$ as $\delta \to 0$. The characteristic function (c.f.) of the limiting distribution is given by

$$\lim_{\delta \to 0} H(e^{\frac{iw_1\delta}{1-\rho}}, e^{\frac{iw_2\delta}{\rho}} | m) = [1 - i(w_1 + w_2)]^{-1} \quad . \tag{55}$$

From this we conclude that as $\delta \to 0$,

$$(\frac{\delta}{1-\rho} X_T, \frac{\delta}{\rho} Z_T) \xrightarrow{\mathcal{L}} (X,Z) \quad , \tag{56}$$

where the c.f. of (X,Z) is given by (55). Furthermore, since w_1 and w_2 appear in (55) as $(w_1 + w_2)$, we have $X = Z$, a.s.. Also the common distribution of X or Z is $\frac{1}{2}\chi^2_2$, so that for small δ, we have the desired approximation

$$X_T \approx \frac{1-\rho}{2\delta}\ \chi^2_2\ ;\ Z_T \approx \frac{\rho}{2\delta}\ \chi^2_2\ . \tag{57}$$

One may, however, question the validity of the approximation (57), since the fact $X = Z$, a.s., yields for small b

$$X_T \approx \frac{1-\rho}{\rho}\ Z_T\ .\ \text{a.s.} \tag{58}$$

Equation (58) implies that if the observed points for (X_T, Z_T) are plotted for various bacterial bursts, they should follow along a straight line rather closely. This, although it cannot be checked experimentally, may very well be false. However, one of the interesting aspects of the above model is the fact that the approximation (57) is independent of m, the number of bacteriophages that infected the bacterium. This fact is compatible with the observations made by Delbrück, where in cases such as virus alpha, he finds no significant difference between the distributions of Z_T for the single and multiple infection cases. This also supports the general findings that the bacteria, which are simultaneously infected with several virus particles of the same kind, react as if only one of the virus particles were effective. (See Delbrück and Luria [3]).

The exponential distribution of Z_T as given by (57) was fitted by the minimum chi-square method to the data on burst size published by Delbrück [2] for the case of single infection with virus alpha. The fit is indicated by the dotted curve A in the figure. The estimate of δ/ρ turns out to be about 0.00415. The calculated value of chi-square (25.4) is found highly significant. Also, the theoretical value

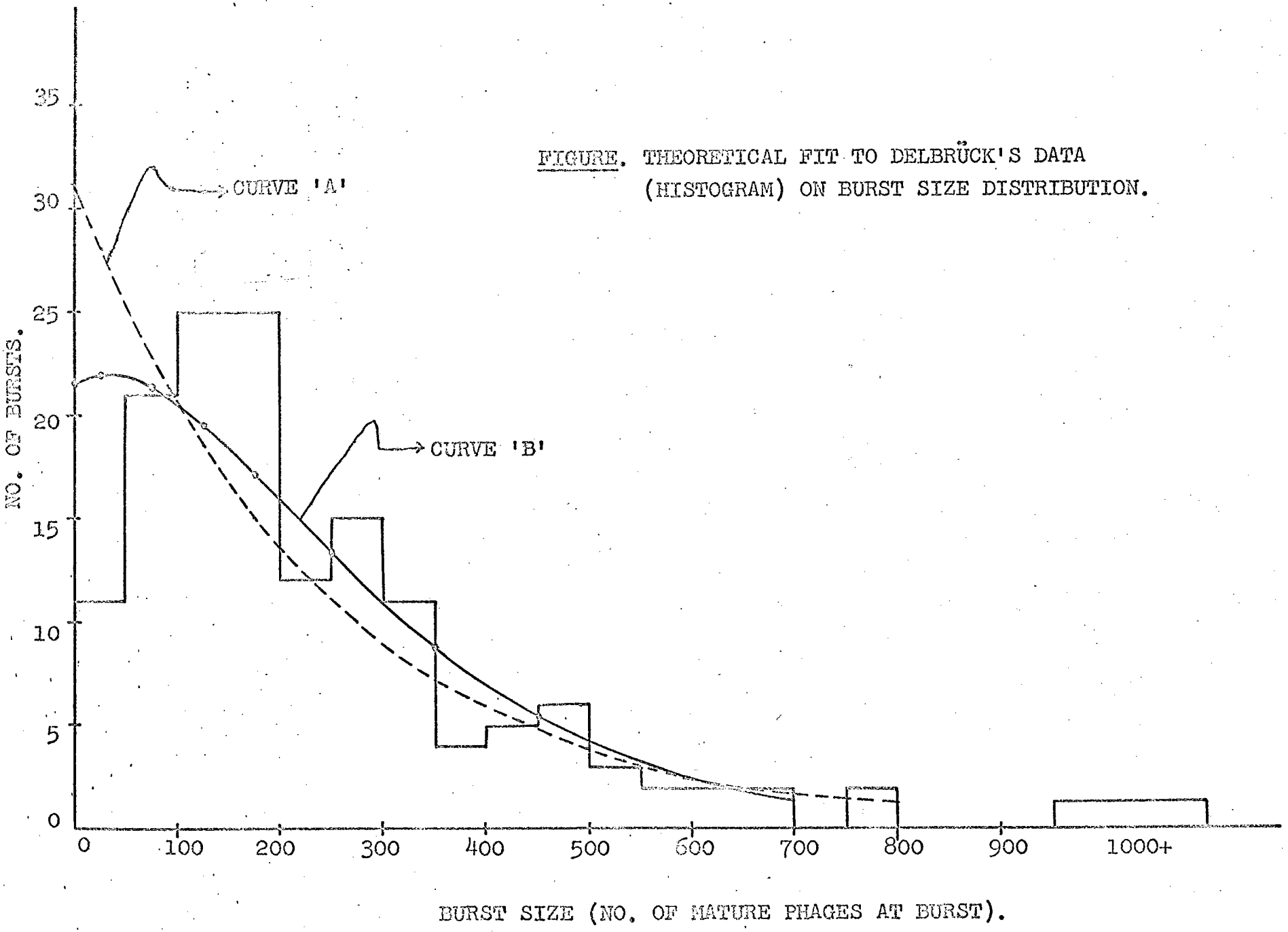

FIGURE. THEORETICAL FIT TO DELBRÜCK'S DATA (HISTOGRAM) ON BURST SIZE DISTRIBUTION.

of the coefficient of variation $[\text{var } Z_T]^{\frac{1}{2}}/[E\ Z_T]$, as obtained by using (57), is 100 percent; whereas its observed value is only 83.3 percent. Thus the distribution of Z_T as it stands fits the observations rather poorly. Several remarks can be made at this point in search for a possible explanation:

(a) It is quite possible that in reality δ is not as small as the above approximation tends to assume, and as such one should try to fit the exact distribution of Z_T given by (51) rather then the one given by (57). This, unfortunately, appears rather cumbersome.

(b) Another possibility is that the simplified model with $\theta = 0$ may be inappropriate, since it undermines the role of the mature bacteriophages in causing the burst. As such, one needs to explore the formula (26) in conjunction with (44). This appears intractable.

(c) Hetrogeneity of bacteria. In the above model, it was tacetly assumed that before infection all the bacteria are homogeneous with respect to all the factors that may have any influence on their burst sizes. However, this may not be true. For instance, it is known that, in general, the bacteria vary in their sizes (volume). They also may be in differnet stages of their growth at the time of infection. Unfortunately, how much these factors influence the burst size distribution is not clear-cut. Further experimentation in this direction and the improvement of the mathematical models based on this, are highly desirable.

(d) An alternative and somewhat convincing explanation comes from a remark made by Delbrück (page 133, [2]), where he states that a certain proportion of the samples which showed bursts, must be ascribed to tubes in which two or more bacteria were lysed. This certainly calls for a modification of (57) to give distribution of the number of mature bacteriophages not for a bacterium but for a tube that contains a nonzero random number M of bacterial bursts. Also, it appears reasonable to assume that M has a poisson distribution with parameter ξ , so that

$$\Pr(M = k \mid M > 0) = \frac{\xi^k e^{-\xi}}{k!(1 - e^{-\xi})} ; \quad k = 1,2,\ldots . \tag{59}$$

Let the random variable $\tilde{Z}$ denote the number of mature bacteriophages in a tube, so that the conditional distribution of $\tilde{Z}$ given $M = k$, is that of $\rho\, \chi^2_{2k}/2\delta$. Here, we have assumed that the bacteria behave independently without influencing their individual burst sizes, so that given $M = k$, $\tilde{Z}$ is the sum of k independent random variables each distributed as $\rho\, \chi^2_2/2\,\delta$. The distribution of $\tilde{Z}$ is thus a mixture of chi-square distributions with truncated Poisson weights. Its density function is given by

$$\varphi_{\tilde{Z}}(z) = \sum_{k=1}^{\infty} \frac{\xi^k e^{-\xi}}{k!(1 - e^{-\xi})} \left(\frac{2\delta}{\rho}\right) f_{\chi^2_{2k}}\left(z\,\frac{2\delta}{\rho}\right) , \tag{60}$$

where $f_{\chi^2_r}(x)$ is the density function of central chi-square with r d.f.. The density function (60) was fitted again by the minimum chi-square method to the data of Delbrück [2]. The fit is exhibited by the curve B in the figure. The estimates of the two parameters ξ and $\left(\frac{\delta}{\rho}\right)$ of (60) are 2.2 and 0.0108 respectively. The calculated value of chi-square (5 d.f.) was about 10.8, which is significant at 0.1 level of significance but not at 0.05 level. We note that the fit has considerably improved although still it is not as satisfactory as one may wish to achieve. Also, the estimate $\hat{\xi} = 2.2$ is considered rather too high for the data predominantly based on single infection. The following possibility appears to give some hope in further improvement of the above model.

(e) Various stages of growth from vegetative to mature phage: Recently, the author's attention was drawn to one of the important factors which the above model has not taken into account, namely, the presence of a number of intermediate stages of growth from vegetative phage to the mature one(see Wood [16]). It is through these stages the assembly of a mature bacteriophage takes place starting with a vegetative phage. The first intermediate stage is the one where the vegetative phage gets covered with a prism -like protein coat that immediately renders it incapable of multiplication. On the other hand, it reaches the final stage of mature bacteriophage only after the attachment of the tail fibres, without which it is not considered capable of infecting a bacterium. Thus a phage in an intermediate stage differs both from vegetative and from mature bacteriophage, in that it neither can multiply nor can infect a bacterium. At the burst of the bacterium, along with the vegetative and mature bacteriophages, a certain random number of phages in the intermediate stages are released. The standard technique of making burst count, which is based on counting plaques on the agar-plate, only picks up the mature bacteriophages, while the phages in the intermediate stages like vegetative phages are considered lost.

The incorporation of the above consideration into the mathematical theory appears essential in order to bring it closer to the reality. As a first approximation it may be sufficient to assume the existence of a single intermediate stage. It is clear, however, that this additional feature of an intermediate stage makes the theory algebraically more involved. Further work in this direction is underway and shall be reported elsewhere in a later communication.

References

[1] Bartlett, M.S. (1961). Equations for stochastic path integrals. Proc. Camb. Phil. Soc., 57, Pt. 3, pp. 568-73.

[2] Delbrück , M. (1945). The burst size distribution in the growth of bacterial viruses (bacteriophages), J. Bacteriology, 50, pp. 131-35.

[3] Delbrück , M. and Luria, S.E. (1942). Interference between bacterial viruses, Arch. Biochem., 1, pp. 111-141.

[4] Gani, J. (1962). An approximate stochastic model for phage reproduction in a bacterium. J. Aust. Math. Soc., 2, pp. 478-483.

[5] Gani, J. (1965). Stochastic models for bacteriophage, J. Appl. Prob., 2, pp. 225-268.

[6] Gani, J. and Yeo, G.F. (1965). Some birth-death and mutation models for phage reproduction, J. Appl. Prob., 2, pp. 150-161.

[7] Kendall, D.G. (1948). On the generalized "Birth-and-Death" process, Ann. Math. Statist., 19, pp. 1 - 15 .

[8] Kimball, A. W. (1965). A model for chemical mutagenesis in bacteriophage, Biometrics, 21, pp. 875-889.

[9] Ohlson, S. (1963). Further models for phage reproduction in a bacterium, Biometrics, 19, pp. 441-449.

[10] Puri, P.S. (1966). On the homogeneous birth-and-death process and its integral, Biometrika, 53, pp. 61-71.

[11] Puri, P. S. (1968). Some further results on the birth-and-death process and its integral, Proc. Camb. Phil. Soc., (In Press).

[12] Puri, P. S. (1967). Some limit theorems on branching processes and certain related processes, Sankhyā. (To appear).

[13] Puri, P. S. (1968). Interconnected birth and death processes, J. Appl. Prob. (In press).

[14] Puri, P. S. (1967). A class of stochastic models of response after infection in the absence of defense mechanism, Proc. 5th Berkeley symposium, 4, pp.511-535.

[15] Steinberg, C. and Stahl, F. (1961). The clone size distribution of mutants arising from a steady-state pool of vegetative phage. J. Theoret. Biol., 1, pp. 488-497.

[16] Wood, W. B. (1967). Gene action in the control of bacteriophage T4 morphogenesis, California Inst. Tech. (unpublished).