

THE REAL POWERS OF THE CONVOLUTION OF A NEGATIVE BINOMIAL DISTRIBUTION AND A BERNOULLI DISTRIBUTION

GÉRARD LETAC, DHAFFER MALOUCHE, AND STEFAN MAURER

(Communicated by Stanley Sawyer)

ABSTRACT. For $c > 0$, this note computes essentially the set of (x, y) in $[0, +\infty)^2$ such that the entire series in z defined by $(1 + z/c)^x(1 - z)^{-y}$ has all its coefficients non-negative. If X and Y are independent random variables which have respectively Bernoulli and negative binomial distributions, denote by μ the distribution of $X + Y$. The above problem is equivalent to finding the set of $p > 0$ such that μ^{*p} exists; this set is a finite union of intervals and may be the first example of this type in the literature. This gives the final touch to the classification of the natural exponential families with variance functions of Babel type, *i.e.* of the form $aR(m) + (bm + c)\sqrt{R(m)}$, where R is a polynomial with degree ≤ 2 .

I. INTRODUCTION: THE BABEL CLASS OF NATURAL EXPONENTIAL FAMILIES

Natural exponential families. If μ is a positive measure on \mathbb{R} , its Laplace transform is the function valued in $(0, +\infty]$ defined by

$$L_\mu(\theta) = \int_{-\infty}^{+\infty} \exp(\theta x) \mu(dx).$$

One considers the interval $\Theta(\mu) =$ interior of the interval of θ such that $L_\mu(\theta)$ is finite. The function $k_\mu = \log L_\mu$ is called the cumulant function of μ , although μ is not necessarily a probability or even bounded. We denote by \mathcal{M} the set of such measures on \mathbb{R} such that, furthermore, μ is not concentrated on a point and such that $\Theta(\mu)$ is not empty. For $\mu \in \mathcal{M}$, the set of probabilities

$$F = F(\mu) = \{P(\theta, \mu)(dx) = \exp(\theta x - k_\mu(\theta))\mu(dx); \theta \in \Theta(\mu)\}$$

is the natural exponential family (NEF) generated by the basis μ (a basis is not unique: $F(\mu) = F(\mu')$ if and only if there exists α and c such that $\mu'(dx) = \exp(\alpha x + c)\mu(dx)$). One also has

$$k'_\mu(\theta) = \int_{-\infty}^{+\infty} x P(\theta, \mu)(dx),$$

Received by the editors May 1, 1998 and, in revised form, November 4, 1998.

1991 *Mathematics Subject Classification.* Primary 60E10; Secondary 33A65.

Key words and phrases. Exponential family, Meixner polynomials, Jorgensen set.

so that the interval $M_F = k'_\mu(\Theta(\mu))$ is called the domain of the means of the NEF F . Since $\mu \in \mathcal{M}$, k_μ is real analytic and strictly convex, the map $\theta \mapsto k'_\mu(\theta)$ is one-to-one, and one defines its inverse function $\psi_\mu : M_F \rightarrow \Theta(\mu)$. Thus

$$m \mapsto P(m, F) = P(\psi_\mu(m), \mu)$$

is a one to one map from M_F onto F , and is called the parametrization of the NEF by the mean.

The variance function of a NEF. The variance of $P(m, F)$ is denoted by $V_F(m)$. The map $m \mapsto V_F(m)$ from M_F to \mathbb{R} is called the variance function of the NEF F . It characterizes F : if F' is another NEF such that $I = M_{F'} \cap M_F$ is not empty and $V_F = V_{F'}$ on I , then $F = F'$; since $k''_\mu(\theta) = V_F(k'_\mu(\theta))$ for all θ in $\Theta(\mu)$, it is easily justified that the data of M_F and V_F characterizes F : see [10] for details. One has also to know that the classical NEF have very simple variance functions: see [10] for a systematic study of the F such that V_F is the restriction to M_F of a polynomial of degree ≤ 2 , and [6] for degree 3.

The Babel class. The next degree of complexity of variance functions after polynomials with degree ≤ 2 or 3 is not offered by polynomials of degree 4, but rather by the class of functions of the form $PR + Q\sqrt{R}$, where the polynomials P , Q and R have degrees ≤ 1 , 2 and 2 respectively. This is called the Grand Babel class, and contains many of the current distributions used in probability theory. Its complete classification is still an open problem (consult [4] and [5]). But the subclass of the variance functions of the form $aR(m) + (bm + c)\sqrt{R(m)}$ is called the Babel class and has been systematically classified in [7], up to a problem that we are going to describe below after a small digression about the real powers of a NEF.

The Jorgensen set of a NEF. Let us come back to an arbitrary element μ of \mathcal{M} . Its Jorgensen set $\Lambda(\mu)$ is the set of $p > 0$ such that there exists $\mu_p \in \mathcal{M}$ with $\Theta(\mu_p) = \Theta(\mu)$ and $L_{\mu_p}(\theta) = (L_\mu(\theta))^p$: see [3]. If μ and μ' are two bases of the NEF F , it is trivial to see that $\Lambda(\mu)$ and $\Lambda(\mu')$ coincide and that similarly $F(\mu_p) = F(\mu'_p)$ (and therefore can be denoted by Λ_F and F_p). Trivially, $\{0\} \cup \Lambda$ is a closed additive semigroup. It contains 1 by definition, therefore it contains the set \mathbf{N} of non-negative integers. It is the whole half line $[0, +\infty)$ if and only if F is made with infinitely divisible distributions. It is \mathbf{N} when F is the family of Bernoulli distributions on $\{0, 1\}$. Finding the Jorgensen set of a given distribution or NEF can be a rather challenging problem: it has been for instance considered implicitly by Paul Lévy ([7]) when k_μ is a polynomial with respect to $\exp \theta$ with some negative coefficients, or by W. Hayman ([2]). Some comments about Paul Lévy's paper can be found in [9], section 8.2: if $P(z) = 1 + z - \frac{z^2}{4} + z^3 + z^4$ and if

$$\exp P(z) = \sum_0^\infty \mu_n z^n,$$

then it is not difficult to see that $\mu_n \geq 0$ for all n and that the Jorgensen set of the discrete measure $\mu(dx) = \sum_0^\infty \mu_n \delta_n(dx)$ is exactly $[1/2, \infty)$. The idea is to observe that P^2 and P^3 have non-negative coefficients and to write

$$\exp pP = pP + \sum_0^{+\infty} p^{2k} \frac{P^{2k}}{(2k)!} + \sum_0^{+\infty} p^{2k+3} \frac{P^3 P^{2k}}{(2k+3)!}.$$

For $p > 0$, the only possible negative coefficient of the entire series is the coefficient $\frac{1}{4}(2p^2 - p)e^p$ of z^2 , and this proves the result.

Similarly, consider the positive measure $\nu(dx) = \sum_0^8 \nu_n \delta_n(dx)$ defined by $P^2(z) = \sum_0^8 \nu_n z^n$ (in fact $\nu_4 = 65/16$, $\nu_3 = \nu_5 = 3/2$, $\nu_2 = \nu_6 = 1/2$, $\nu_1 = \nu_7 = 2$, $\nu_0 = \nu_8 = 1$). It provides the non classical Jorgensen set

$$\{1, 3/2, 2, 5/2, 3, \dots\}.$$

The ultimate problem. In [7], page 98, it was left as an open problem to find the Jorgensen set of the measure on \mathbf{N} :

$$\mu(dx) = \sum_{n=0}^{\infty} \mu(n) \delta_n(dx)$$

where the sequence $(\mu(n))_{n=0}^{\infty}$ is defined by

$$\sum_{n=0}^{\infty} \mu(n) z^n = \frac{1 + ze^{-\alpha}}{(1 - ze^{\alpha})^q},$$

for fixed $q > 0$ and fixed $\alpha \in \mathbb{R}$. Here, $F = F(\mu)$ is the NEF of the probability distributions of the random variables $X + Y$, where the independent random variables have the following distributions

$$\Pr(X = 0) = 1 - r, \quad \Pr(X = 1) = r,$$

$$\Pr(Y = n) = \frac{q(q + 1) \cdots (q + n - 1)}{n!} \lambda^n (1 - \lambda)^q, \quad n \in \mathbf{N};$$

the parameters λ and r in $(0, 1)$ are arbitrary, subject to the constraint $r \exp(2\alpha) = (1 - r)\lambda$. Its variance function is in the Babel class. One has $M_F = (0, \infty)$ and, by a tedious calculation, or by using formula (3.3), page 83, of [7]:

$$V_F(m) = \frac{1}{4q \cosh^2 \alpha} ((q - 1)R(m) + (4m + 2(1 - c))\sqrt{R(m)})$$

with

$$c = \frac{1}{2}(e^{-\alpha} + qe^{\alpha}), \quad R(m) = m^2 \cosh^2 \alpha + m(q - 1 + 2c \sinh \alpha) + c^2.$$

The set of $p > 0$ such that $V_{F_p}(m) = pV_F(\frac{m}{p})$ is a variance function, or equivalently, such that the Taylor expansion of

$$(1.1) \quad z \mapsto \frac{(1 + ze^{-\alpha})^p}{(1 - ze^{\alpha})^{pq}}$$

has positive coefficients, is going to be described in the present note. For $\alpha < 0$ a simple argument in Section 2 will show that this set is exactly \mathbf{N} . For $\alpha \geq 0$ the answer is more surprising and is described in Section 3.

For convenience, in the function (1.1) we are going to change the notations used in [7] and to replace p by x , pq by y , $\exp 2\alpha$ by c and $z \exp \alpha$ by z . Thus, for fixed $c > 0$, we consider the sequence $(a_n(x, y, c))_{n=0}^{\infty}$ defined by

$$(1.2) \quad \frac{(1 + z/c)^x}{(1 - z)^y} = \sum_{n=0}^{\infty} a_n(x, y, c) z^n,$$

and the closed additive semigroup of $[0, +\infty)^2$ defined by

$$(1.3) \quad S_c = \{(x, y); x \geq 0, y \geq 0, a_n(x, y, c) \geq 0 \forall n \in \mathbf{N}\}.$$

The choice of the parameter c is made to facilitate the use of a traditional notation of the Meixner polynomials that we are going to use in Section III.

Our problem is now to describe S_c . Let us make some simple remarks about it. Trivially $\mathbf{N} \times \{0\}$ and $\{0\} \times [0, +\infty)$ are subsets of S_c . This implies in particular that given $x \geq 0$, if there exists $y \geq 0$ such that $(x, y) \in S_c$, then $(x, y') \in S_c$ for $y' > y$. Thus if we consider the function $f_c : [0, +\infty) \rightarrow [0, +\infty]$ defined by

$$(1.4) \quad f_c(x) = \inf\{y; (x, y) \in S_c\},$$

then the description of S_c is reduced to the description of the function f_c (note that it can take the value $+\infty$), since

$$S_c = \{(x, y); x \geq 0, y \geq 0, y \geq f_c(x)\}$$

and since S_c is a closed set.

It may be worth mentioning that the present paper originated in an attempt to improve the following majorization of f_1 by a function of period 1:

$$\text{for } k - 1 \leq x \leq k - \frac{1}{2}, \quad f_1(x) \leq x - k + 1;$$

$$\text{for } k - \frac{1}{2} \leq x \leq k, \quad f_1(x) \leq k - x.$$

They are proved respectively by writing

$$\frac{(1+z)^x}{(1-z)^{x-k+1}} = (1+z)^{k-1} \exp[(x-k+1) \sum_0^{\infty} \frac{2}{2n+1} z^{2n+1}];$$

$$\frac{(1+z)^x}{(1-z)^{k-x}} = \frac{(1+z)^k}{(1-z^2)^{k-x}}.$$

II. THE CASE $0 < c < 1$

Theorem 2.1. *For $0 < c < 1$ and $y > 0$, the Taylor expansion of*

$$z \mapsto \frac{(1+z/c)^x}{(1-z)^y}$$

has non-negative coefficients if and only if x is in the set \mathbf{N} of non-negative integers.

Proof. The “if” part is obvious. We offer two proofs of the “only if” part.

1) Assume that there exists $x > 0$ with $x \notin \mathbf{N}$ such that the entire series (1.2) has positive coefficients. Trivially, its radius of convergence is c . Consider the positive measure on \mathbf{N} :

$$\mu = \sum_{n=0}^{\infty} a_n(x, y, c) \delta_n.$$

With the notations of the introduction, one has $\Theta(\mu) = (-\infty, \log c)$. However, the Laplace transform L_μ is

$$L_\mu(\theta) = \frac{(1 + e^\theta/c)^x}{(1 - e^\theta)^y}$$

on $\Theta(\mu)$ and is analytically extendable by a positive function to the bigger interval $(-\infty, 0)$. From a well known result of [12], this is impossible and μ cannot be positive. The theorem is proved.

2) By Darboux’s method ([11], Theorem 8.4) we have the asymptotic relation

$$a_n(x, y, c) \sim \left(1 + \frac{1}{c}\right)^x \binom{-y}{n} + (1+c)^y \binom{x}{n} c^{-n}.$$

This shows that for $0 < c < 1$ and $y > 0$, the a_n alternate in sign for n sufficiently large, proving the theorem.

III. THE CASE $c \geq 1$

Here is our main result, that we state with the notations (1.2)–(1.4).

Theorem 3.1. *Assume $c \geq 1$. Let $k - 1 < x < k$, where k is a positive integer. Then (x, y) is in S_c if and only if $a_{k+1}(x, y, c) \geq 0$.*

Proof of Theorem 3.1. As usual we consider the polynomial of degree k :

$$\binom{x}{k} = \frac{x(x-1) \cdots (x-k+1)}{k!}$$

and we write explicitly

$$(3.1) \quad a_n(x, y, c) = \sum_{l=0}^n c^{-l} \binom{x}{l} (-1)^{n-l} \binom{-y}{n-l}.$$

As a consequence:

$$(3.2) \quad a_n(x, y, c) > 0 \text{ if } x > n - 1.$$

Let us now fix $x \in (k-1, k)$. Inequality (3.2) implies the nonnegativity of a_1, a_2, \dots, a_k .

The essential fact is now that the a_n satisfy the recursion:

$$(3.3) \quad c(n+2)a_{n+2} = [x + cy + (c-1)(n+1)]a_{n+1} + [y + (n-x)]a_n$$

as follows from the differential equation

$$c(1-z)(1+z/c) \frac{d}{dz} \frac{(1+z/c)^x}{(1-z)^y} = [x + cy - (x-y)z] \frac{(1+z/c)^x}{(1-z)^y}.$$

To finish the proof, assume that $a_{k+1} \geq 0$. Then using $n = k, k+1, \dots$ in (3.3) shows by induction that $a_n \geq 0$ for $n > k$.

Comments: 1) Recall that $\mathbf{N} \times [0, +\infty) \subset S_c$ so there is nothing to prove if x is an integer.

2) For fixed x in $(k-1, k)$, then $y \mapsto a_{k+1}(x, y, c)$ is a polynomial. The theorem shows that its smallest non-negative root is exactly $f_c(x)$. More precisely, from (3.1) one has trivially $a_{k+1}(x, 0, c) < 0$, the coefficient of highest degree of $a_{k+1}(x, y, c)$ is positive, thus $f_c(x)$ exists and is strictly positive. The set of non-negative y such that $a_{k+1}(x, y, c) < 0$ must be an interval, which is $[0, f_c(x))$. Finally, $f_c(x)$ is the only root of $a_{k+1}(x, y, c)$ in $[0, \infty)$. This comes from the fact that $y \mapsto a_{k+1}(x, y, c)$ is a non decreasing function on $[f_c(x), \infty)$. To see this, observe that $f_c(x) \leq y \leq y'$ implies

$$a_{k+1}(x, y', c) = \sum_{l=0}^{k+1} a_l(0, y' - y, c) a_{k+1-l}(x, y, c) \geq a_{k+1}(x, y, c)$$

since $a_0 = 1$ and since $a_l(0, y' - y, c)$ and $a_{k+1-l}(x, y, c)$ are non-negative for $l = 0, 1, \dots, k$.

3) For instance, to get the graph of f_c on $[0, 1]$, one computes

$$2a_2(x, y, c) = y^2 + \left(1 + \frac{2x}{c}\right)y + \frac{x^2 - x}{c^2},$$

which has one positive root

$$f_c(x) = -\frac{x}{c} - \frac{1}{2} + \sqrt{\frac{x(1+c)}{c^2} + \frac{1}{4}}.$$

Its graph is obtained by drawing in the (x, y) plane the parabola defined by $a_2(x, y, c) = 0$. For higher values of k , it is difficult to be explicit about the values of f_c . A computation with MAPLE provides the following graph for f_1 :

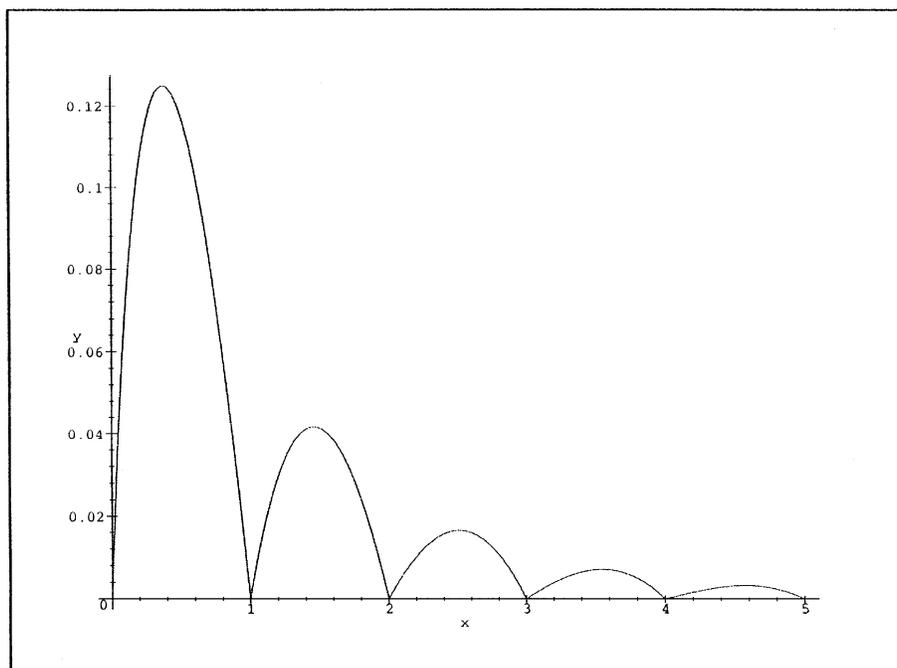


FIGURE 1.

4) This function f_c is analytic in $(k-1, k)$, since it is not difficult to prove that $\frac{\partial}{\partial y} a_{k+1}(x, y) > 0$ for $(x, y) \in (k-1, k) \times (0, +\infty)$. Applying the implicit function theorem to the polynomial a_{k+1} and to the domain $(k-1, k) \times (0, +\infty)$ gives the result.

5) Consider the function g_k defined on $[0, 1]$ by $g_k(x) = f_c(k-1+x)$ where k is a positive integer. Since S_c is a semigroup which contains $(1, 0)$, this implies that $(g_k(x))_{k \geq 1}$ is a decreasing sequence. A more detailed study can show for instance that for $c = 1$ one has

$$\max\{g_k(x); 0 \leq x \leq 1\} \sim_{k \rightarrow \infty} \frac{1}{2e} \frac{1}{2^k \log k}.$$

It is not difficult to compute

$$g'_k(0) = \frac{1}{1 + (ck-1)(c+1)^k} \quad \text{and} \quad g'_k(1) = \frac{1}{1 - (c+1)^k}.$$

Although Figure 1 suggests that g_k are concave on $[0, 1]$, this is not true. Using MAPLE one finds for $c = 1$ and for $k = 5, 6, 7$ that $g_k''(0) > 0$. One can even reasonably conjecture that g_k has one inflection point in $[0, 1]$ whose limit is 1 when k goes to infinity.

6) For $q > 0$, $(1, q)$ is in S_c ; let us consider the positive measure

$$\mu = \sum_{n=0}^{\infty} a_n(1, q, c)\delta_n.$$

Its Jorgensen set is the set of $p > 0$ such that (p, pq) belongs to S_c and is visualized by drawing the line in the plane through the origin and the point $(1, q)$: one retains the part of the line contained in S_c . If the slope of the line is small enough, one gets an explicit example of a Jorgensen set Λ which is neither $(0, \infty)$ nor $\lambda\mathbf{N}$, but is made with the union of small intervals completed by a half line. Analysing the roots x of the two equations $a_2(x, qx, c) = 0$ and $a_3(x, qx, c) = 0$, one gets the following partial description of Λ :

If $1 - qc^2 \geq 0$, then $\Lambda = (0, +\infty)$.

If $q_c \leq q < 1/c^2$, where q_c is the smallest root of $c^2q^2 - 2(8c^2 - 9c + 8)cq + 1 = 0$, then

$$\Lambda = [(1 - qc^2)/(1 + qc), +\infty).$$

If $q < q_c$, and if $1 < x_- < x_+ < 2$ are the roots of

$$(1 + cq)^3x^2 - 3(1 + cq)(1 + cq^2)x + 2(1 + c^3q) = 0,$$

then

$$\Lambda = [(1 - qc^2)/(1 + qc), x_-] \cup \Lambda_1,$$

where $x_+ \in \Lambda_1 \subset [x_+, \infty)$.

7) To put recursion (3.3) in perspective, we note that the Meixner polynomials $m_n(x; \beta; -c)$ satisfy

$$\frac{(1 + z/c)^x}{(1 - z)^{x+\beta}} = \sum_0^{\infty} \frac{m_n(x; \beta; -c)}{n!} z^n.$$

They can be expressed through the hypergeometric function

$$m_n(x; \beta; -c) = (\beta + x)_n F(-n, -x; 1 - \beta - x; -c^{-1}) = (\beta)_n F(-n, x; \beta; 1 + c^{-1}).$$

They are related to our a_n by $a_n(x, y, c) = \frac{1}{n!} m_n(x; y - x; -c)$. In fact, (3.3) follows from the well-known three-term recursion

$$cm_{n+1} = [(c + 1)x + (c - 1)n + c\beta]m_n + n(n + \beta - 1)m_{n-1},$$

a special case of Gauss' contiguous relation for the hypergeometric function (see [1]).

REFERENCES

[1] ERDELYI, A., MAGNUS, W., OBERHETTINGER, F., TRICOMI, F. G., et al (1953), *Higher Transcendental Functions. Vol. 2.* New York, McGraw-Hill Book Co. MR **15**:419i
 [2] HAYMAN, W. K. (1956), "A generalisation of Stirling's formula." *J. Reine Angew. Math.*, **196**, 67-95. MR **18**:293f
 [3] JORGENSEN, B. (1987), "Exponential dispersion models." *J. Roy. Statist. Soc. Ser. B*, **49**, 127-162. MR **88m**:62094
 [4] KOKONENDJI, C. C. (1994), "Exponential families with variance functions in $\sqrt{\Delta}P(\sqrt{\Delta})$: Seshadri's class." *Test* **3** 123-172. MR **97c**:62033

- [5] KOKONENDJI, C. C. (1995), “Sur les familles exponentielles naturelles de grand-Babel.” *Annales de la Faculté des Sciences de Toulouse, IV,4*, 763- 799. MR **99f**:62024
- [6] LETAC, G. and MORA, M. (1990), “Natural real exponential families with cubic variances.” *Ann. Statist.* *18*, 1-37. MR **91b**:62032
- [7] LETAC, G. (1992), *Lectures on Natural Exponential Families and their Variance Functions*. Monografias de Matemática, Vol 50. Rio de Janeiro: Instituto de Matemática Pura e Aplicada. MR **94f**:60020
- [8] LÉVY, P. (1937), “Sur les exponentielles de polynômes.” *Annales de l'Ecole Normale Supérieure*, *73*, 231-292.
- [9] LUKACS, E. (1970), *Characteristic Functions*. 2nd edition. Charles Griffin and Co., London. MR **49**:11595
- [10] MORRIS, C. N. (1982), “Natural exponential families with quadratic variance functions.” *Ann. Statist.* *10*, 65-80. MR **83a**:62037
- [11] SZEGÖ, G. (1975), *Orthogonal Polynomials* (4 th edition) AMS, Providence, R.I. MR **51**:8724
- [12] WIDDER D. V. (1941), *The Laplace Transform*. Princeton University Press, New Jersey. MR **3**:232d

LABORATOIRE DE STATISTIQUE ET PROBABILITÉS, UNIVERSITÉ PAUL SABATIER, 31062, TOULOUSE, FRANCE

E-mail address: `letac@cict.fr`

24 AV. MONGI SLIM, 1004 EL MENZAH V, TUNISIE

E-mail address: `dhafer_malouche@yahoo.fr`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF VIRGINIA, CHARLOTTESVILLE, VIRGINIA 22902

E-mail address: `srm4x@virginia.edu`