Søren Johansen

# Homomorphisms and General Exponential Families

Preprint
September
1976

14

Institute of Mathematical Statistics
University of Copenhagen

Søren Johansen

# HOMOMORPHISMS AND GENERAL EXPONENTIAL FAMILIES*

Preprint 1976 No. 14

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN
and
DEPARTMENT OF STATISTICS
OREGON STATE UNIVERSITY

September 1976

* Research for this paper was supported by the National Science Foundation, under Grant MCS 74-07589 A01 with Oregon State University.

KØBENHAVNS UNIVERSITET
INSTITUT FOR
MATEMATISK STATISTIK

Søren Johansen
1976

# Homomorphisms and general exponential families

## 1 Introduction

An exponential type family is usually defined by its densities

$$e^{\alpha(\theta)'t(x) + B(x) + C(\theta)}$$

with respect to some measure $\mu_o$. Here $\alpha(\theta) \in R^r$, $\theta \in \Theta$ and $t(x) \in R^r$, $x \in X$. Clearly the factor $B(x)$ can be taken into $\mu_o$ and $C(\theta)$ is just a normalizing constant. Thus the interplay between the observation and the parameter is given by the bilinear form

$$\alpha(\theta)'t(x). \tag{1.1}$$

In this formulation the parameter $\alpha(\theta)$ and the observation $t(x)$ have a dual role, in that $t(x)$ can be considered a linear functional on the range of $\alpha(\theta)$ and vice versa [2].

Of special interest are the models obtained by allowing $\alpha(\theta)$ to vary freely in the set

$$D = \{\alpha \mid \phi(\alpha) = \int e^{\alpha' t(x)}\, d\mu < \infty\}$$

and we then get a canonical (or full) model defined by its densities

$$\frac{e^{\alpha' t(x)}}{\phi(\alpha)}, \quad \alpha \in D$$

with respect to $d\mu = e^{B(x)}\, d\mu_o$.

The purpose of this note is to show how in any family of probability measures, a similar duality holds. The basic idea is to represent x by its likelihood function and $\theta$ by the density and show how this represents $(x,\theta)$ as a point in a pair of semigroups in duality thus giving a general formulation of (1.1). The idea of using semigroups and homomorphisms as tools in the discussion of exponential families is due to Lauritzen [3].

Once this formulation is given one has the possibility of defining a general concept of a canonical family and to show how one can extend a given family to a canonical family.

The report presents part of the work done jointly with H.D.Brunk, D.Birkes and J.Lee at Oregon State University, Corvallis, see also [1].

## 2 The formulation

To avoid measure theoretic difficulties the following set up will be used: $X$ is a separable topological space and $P = \{P_\theta, \theta \in \Theta\}$ a family of probabilities on $X$. Let $S_\theta$ denote the support of $P_\theta$ and $\mu$ a $\sigma$-finite measure on $P$. We assume that $P_\theta$ has density $f(x,\theta)$ with respect to $\mu$, where $\mu$ is equivalent to $P$ and that $f(x,\theta) = 0$, $x \notin S_\theta$ and $f(x,\theta)$ is continuous on $S_\theta$. This defines the density uniquely.

Let $S(X)$ denote the semigroup of finite ordered samples from $X$, i.e. the set of samples $(x_1, \dots, x_n)$ with composition given by

$$(x_1, \dots, x_n)(y_1, \dots, y_m) = (x_1, \dots, x_n, y_1, \dots, y_m).$$

On the space $S(X)$ we define two equivalence relations. For $x = (x_1, \dots, x_n) \in S(X)$ and $y = (y_1, \dots, y_m) \in S(X)$ we define

$$(2.1) \qquad x \sim y \text{ if } \prod_{i=1}^{n} f(x_i,\theta) = \prod_{j=1}^{m} f(y_j,\theta), \ \theta \in \Theta$$

Thus two samples are equivalent if they have the same (strict) likelihood function.

Similarly we define

$$(2.2) \qquad x \approx y \text{ if } \exists\, \lambda > 0 : \prod_{i=1}^{n} f(x_i,\theta) = \lambda \prod_{j=1}^{m} f(y_j,\theta), \ \theta \in \Theta.$$

Thus $x \approx y$ if $x$ and $y$ have the same likelihood function.

We shall also write for $\lambda > 0$

$$(2.3) \qquad [x] = \lambda[y] \quad \text{if} \quad \prod_{i=1}^{n} f(x_i,\theta) = \lambda \prod_{j=1}^{m} f(y_j,\theta)$$

and we thus have

$$x \approx y \quad \text{iff} \quad \exists\, \lambda > 0 : \ [x] = \lambda[y].$$

Now we have the natural mappings

$$X \xrightarrow{t_1} S(X) \xrightarrow{t_2} S(X)/\sim \xrightarrow{t_3} S(X)/\approx,$$

where x is mapped into the one-element sample, and $t_2$ and $t_3$ maps elements into equivalence classes.

Notice that the mapping $t = t_3 \circ t_2$ is the mapping that to each sample associates its likelihood function. Thus t is minimally sufficient in the sense that it induces the minimal sufficient σ-algebra, see Loève [4].

Similarly we consider the parameter space Θ and form S(Θ) and define, for $\tau = (\tau_1, \ldots, \tau_n) \in S(\Theta)$ and $\sigma = (\sigma_1, \ldots, \sigma_m) \in S(\Theta)$ the equivalence relation

$$\tau \sim \sigma \Leftrightarrow \prod_{i=1}^{n} f(x,\tau_i) = \prod_{j=1}^{m} f(x,\sigma_j), \; x \in X. \tag{2.4}$$

Thus in particular $\theta_1$ and $\theta_2$ (in Θ) are equivalent if they correspond to the same density. Notice that since proportional densities are equal (they integrate to 1) there is no equivalent to ≈ on the parameterspace.

Note also that $\sigma = (\sigma_1, \ldots, \sigma_n)$, $\sigma_i \in \Theta$, $i=1, \ldots, n$ corresponds to a measure whose density is a product of densities in the family $P$.

We then have the natural mapping

$$\Theta \xrightarrow{\alpha_1} S(\Theta) \xrightarrow{\alpha_2} S(\Theta)/\sim$$

where the parameter $\alpha_2(\alpha_1(\theta))$ is the maximal identifiable parameter.

We can then formulate the main result as:

Proposition 1 The semigroups $S(X)/\sim$ and $S(\Theta)/\sim$ are in duality by the bihomomorphism $\circ$ defined by

$$[x] \circ [\sigma] = \prod_{i=1}^{n} \prod_{j=1}^{m} f(x_i,\sigma_j). \tag{2.5}$$

Proof Note that [] denotes the equivalence class corresponding to $\sim$, thus $[x] = t_2(x)$ and $[\sigma] = \alpha_2(\sigma)$.

Let us first remark that by the definition of $\sim$, the right hand side of (2.5) only depends on the equivalence class of x and $\sigma$.

It is easily seen that $\circ$ is a bihomomorphism, since for instance,

$$[x] \circ [\sigma\tau] = \prod_{i=1}^{n} \left[ \prod_{j=1}^{m} f(x_i,\sigma_j) \prod_{k=1}^{r} f(x_i,\tau_k) \right]$$

$$= \prod_{i=1}^{n} \prod_{j=1}^{m} f(x_i,\sigma_j) \prod_{i=1}^{n} \prod_{k=1}^{r} f(x_i,\tau_k)$$

$$= ([x] \circ [\sigma])([x] \circ [\tau]).$$

If

$$[x] \circ [\tau] = [x] \circ [\sigma], \quad x \in S(X)$$

then in particular for $x \in X$ we get

$$\prod_{k=1}^{r} f(x,\tau_k) = \prod_{j=1}^{m} f(x,\sigma_j), \quad x \in X$$

which by definition means that $\tau \sim \sigma$ and hence $[\tau] = [\sigma]$. Thus the bihomomorphism separates points and the semigroups are in duality.

The content of (2.5) is that it generalizes (1.1) which holds for ordinary exponential families. The following examples will illustrate the formulation, but the essence of (2.5) is that the parameter can be viewed as a homomorphism on the space of (strict) likelihood functions, and the observation as a homomorphism on the space of densities.

## 3 Some examples

Example 1 Let the densities be given by

$$f(x,\theta) = \frac{e^{\alpha(\theta)'t(x)}h(x)}{\phi(\theta)} \tag{3.1}$$

with respect to Lebesgue measure $\mu$ on $R$. Assume further that t and h are continuous and that $h > 0$.

In order that the essential features of the approach be apparent we will also assume that the functions $\{1, \alpha(\theta), \theta \in \Theta\}$ are linearly independent, and that the functions $\{1, \ln h(x), t(x), x \in R\}$ are linearly independent.

In this case $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_m)$ are equivalent, $x \sim y$, if

$$\alpha(\theta)'\left(\sum_{i=1}^{n} t(x_i) - \sum_{j=1}^{m} t(y_j)\right) + \sum_{i=1}^{n} \ln h(x_i) - \sum_{j=1}^{m} \ln h(y_j) - (n-m)\ln\phi(\theta) = 0,$$

$\theta \in \Theta$.

Now since $\ln \phi(\theta)$ is strictly convex as a function of $\alpha(\theta)$, we must have $n = m$. By the independence of the vectors $\{1, \alpha(\theta), \theta \in \Theta\}$ it follows that

$$\sum_{i=1}^{n} t(x_i) = \sum_{j=1}^{m} t(y_j)$$

and

$$\sum_{i=1}^{n} \ln h(x_i) = \sum_{j=1}^{m} \ln h(y_j).$$

Hence the equivalence relation $\sim$ on $S(X)$ is induced by the mapping

$$t_2((x_1, \dots ,x_n)) = (n, \sum_{i=1}^{n} t(x_i), \sum_{i=1}^{n} \ln h(x_i)),$$

and $S(X)/\sim$ can be represented by a certain subsemigroup of

$$S = (N,+) \times (R^r,+) \times (R,+).$$

Similarly $\sigma = (\sigma_1, \dots ,\sigma_n)$ is equivalent to $\tau = (\tau_1, \dots ,\tau_m)$ if

$$\left(\sum_{i=1}^{n} \alpha(\sigma_i) - \sum_{j=1}^{m} \alpha(\tau_j)\right)' t(x) + (n - m)\ln h(x) - \sum_{i=1}^{n} \ln \phi(\sigma_i) + \sum_{j=1}^{m} \ln \phi(\tau_j) = 0,$$

$x \in X$.

By the independence of $\{1, \ln h(x), t(x), x \in X\}$ it follows, that

$$n = m,$$

$$\sum_{i=1}^{n} \alpha(\sigma_i) = \sum_{j=1}^{m} \alpha(\tau_j)$$

$$\sum_{i=1}^{n} \ln \phi(\sigma_i) = \sum_{j=1}^{m} \ln \phi(\tau_j)$$

and hence that the equivalence $\sim$ on the parameterspace is given by the function

$$\alpha_2(\sigma_1, \dots ,\sigma_n) = \left(- \sum_{i=1}^{n} \ln \phi(\sigma_i), \sum_{i=1}^{n} \alpha(\sigma_i), n\right)$$

and that the semigroup $S(\Theta)/\sim$ can be represented by a certain subsemigroup of

$$(R,+) \times (R^r,+) \times (N,+)$$

The parameter $\theta$ can be identified with homomorphism given by the coefficients

$$(-\ln \phi(\theta),\ \alpha(\theta),\ 1)$$

Notice that not all homomorphisms on S can occur, since $\phi(\theta)$ is determined from $\alpha(\theta)$ by normalization and 1 occurs as the last coefficient. Notice also that the semigroup $S(\theta)/\sim$ as well as the bihomomorphism between $S(\theta)/\sim$ and $S(X)/\sim$ depend on h, the underlying measure.

Now consider (2.3). In a similar way we find that $[x] = \lambda[y]$ iff

$$n = m$$

$$\sum_{i=1}^{n} t(x_i) = \sum_{j=1}^{m} t(y_j) \tag{3.2}$$

$$\ln \lambda = \sum_{i=1}^{n} \ln h(x_i) - \sum_{j=1}^{m} \ln h(y_j)$$

and hence $x \approx y$ iff $n = m$ and $\sum_{i=1}^{n} t(x_i) = \sum_{j=1}^{m} t(y_j)$, thus the equivalence relation is induced by the function

$$s(x_1, \ldots, x_n) = \left(n, \sum_{i=1}^{n} t(x_i)\right)$$

which is known to be minimally sufficient, and which is independent of the underlying measure.

<u>Example 2</u> Let the densities be given by

$$f(x,\theta) = \begin{cases} \frac{1}{\theta} & , \quad 0 \leq x \leq \theta, \\ 0 & , \quad \theta < x < \infty, \end{cases} \tag{3.3}$$

where $0 < \theta < \infty$. Then

$$(x_1, \ldots, x_n) \sim (y_1, \ldots, y_m) \quad \text{iff}$$

$$\frac{1}{\theta^n} 1_{[0,\theta]}(x_{(m)}) = \frac{1}{\theta^m} 1_{[0,\theta]}(y_{(m)}), \quad 0 < \theta < \infty,$$

which happens only if $n = m$ and $x_{(n)} = y_{(m)}$. Thus $S(X)/\sim$ can be identified with $(N,+) \times (R_+, \vee)$ and $\theta$ with the homomorphism $h(n,x) = h_1(n)\, h_2(x)$ where

$$h_1(n) = \theta^{-n}$$

$$h_2(x) = 1_{[0,\theta]}(x).$$

Notice that in this case $x \approx y$ iff $x \sim y$ and this corresponds to $(n, x_{(n)})$ being minimally sufficient.

<u>Example 3</u> Let $P_1$ and $P_2$ be given by two positive continuous densities $f_1$ and $f_2$ on $R$, with respect to Lebesgue measure. Then $(x_1, \ldots, x_n) \sim (y_1, \ldots, y_m)$ when

$$\prod_{i=1}^{n} f_k(x_i) = \prod_{j=1}^{m} f_k(y_j), \quad k = 1,2$$

i.e. when

$$t_1(x_1, \ldots, x_n) = \left(\prod_{i=1}^{n} f_1(x_i), \prod_{i=1}^{n} f_2(x_i)\right)$$

takes on the same value at x and y. Thus $S(X)/\sim$ is a subset of $(R_+, \cdot)^2$. Further $[x] = \lambda[y]$ if

$$\prod_{i=1}^{n} f_k(x_i) = \lambda \prod_{j=1}^{m} f_k(y_j), \quad k = 1,2$$

which occurs if

$$\frac{\prod_{i=1}^{n} f_1(x_i)}{\prod_{i=1}^{n} f_2(x_i)} = \frac{\prod_{j=1}^{m} f_1(y_j)}{\prod_{j=1}^{m} f_2(y_j)}$$

and then $\lambda$ is given by

$$\lambda = \frac{\prod_{i=1}^{n} f_1(x_i)}{\prod_{j=1}^{m} f_1(y_j)} .$$

Thus the equivalence relation $\approx$ is induced by the likelihood ratio

$$t_2(x_1, \ldots, x_n) = \frac{\prod_{i=1}^{n} f_1(x_i)}{\prod_{i=1}^{n} f_2(x_i)}$$

which is known to be minimally sufficient.

## 4.1 The canonical family

In order to obtain the proper definition of a canonical family we shall need the following mathematical object: a semigroup S with scalar multipliers i.e. there exist an operation o an S such that

$$x \circ (y \circ z) = (x \circ y) \circ z$$

and an mapping $*$ on $\mathcal{R}_+ \times S \to S$ such that

$$\lambda * (x \circ y) = (\lambda * x) \circ y = x \circ (\lambda * y)$$

We let S* denote the set of homogeneous homomorphisms on S, i.e. mappings $h: S \to (\mathcal{R}_+, \cdot)$ such that

$$h(x \circ y) = h(x)\ h(y)$$

$$h(\lambda * x) = \lambda\ h(x)$$

Now let $(\mathcal{X},\mu)$ be a measure space and $t: \mathcal{X} \to S$ a statistic

Definition 4.1 A canonical family determined by $(\mu, t, S)$ is given by densities with respect to $\mu$ of the form

$$(4.1) \qquad \frac{\theta(t(x))}{\int \theta(t(x))\ \mu(dx)}, \quad \theta \in D,$$

where

$$D = \{\theta \in S^* \mid \theta(t(x)) \text{ is measurable and } 0 < \int \theta(t(x))\ \mu(dx) < \infty\}.$$

The reason that we use only the homogeneous homomorphisms is that we want the family (4.1) to be essentially independent of $\mu$, in the following sense:

Proposition 4.2 Let $\mu_o$ be equivalent to $\mu$ and $h = \frac{d\mu}{d\mu_o}$. Let $t_o(x) = h(x) * t(x)$ then the families determined by $(\mu, t, S)$ and $(\mu_o, t_o, S)$ are the same.

Proof  This follows since

$$\theta(t_o(x))\, d\mu_o = \theta(h(x) * t(x))\, d\mu_o = h(x)\, \theta(t(x))\, d\mu_o = \theta(t(x))\, d\mu.$$

Thus the densities are the same and the two families are identical.

We now want to show how the family $P$ considered in section 2 can be extended to a canonical family and finally give some examples how this extension works.

We then start with $P$ given by the densities $f(x,\theta)$ on $X$ with respect to $\mu$.

First choose S as the space of multiples of likelihood functions

$$S = \{\lambda \prod_{i=1}^{n} f(x_i,\cdot),\ \lambda > 0,\ i = 1, \ldots ,n\}$$

This is a semigroup with scalar multipliers. To each $x = (x_1, \ldots ,x_n)$ we associate as before the (strict) likelihood function i.e. $t(x) = [x]$.

Thus we are generating S from $S(X)/\sim$ by adding all multiples of elements of $S(X)/\sim$.

Now consider the family $\tilde{P}$ generated by $(\mu, t, S)$. The mapping $\theta \to [\theta]$ associates with each $\theta \in \Theta$ a homogeneous homomorphism on S

$$[\theta] \circ (\lambda * [x]) = \lambda[\theta] \circ [x]$$

since both sides equal $\lambda \prod_{i=1}^{n} f(x_i,\theta)$, $x = (x_1, \ldots ,x_n)$.

The density corresponding to $[\theta]$ is given by $[\theta] \circ [t(x)] = f(x,\theta)$, $x \in X$ and the family $P$ is thus imbedded into $\tilde{P}$.

Corollary 4.3 The family $\tilde{P}$ does not depend on the choice of $\mu$.

Proof Let $\mu_o$ be another measure equivalent to $P$ and let $h = \frac{d\mu}{d\mu_o}$. Likelihoods from densities with respect to $\mu_o$ are proportional to likelihoods from $\mu$ and hence the mapping

$$t_o(x) = [x]_{\mu_o}$$

i.e. the likelihood with respect to $\mu_o$, sends $X$ into S. Note, however, that $t_o(x) = h(x) * [x]_\mu$, where $h(x) = \prod_{i=1}^{n} h(x_i)$ and hence $t_o = h * t$ and it follows from Proposition 4.2 that the families generated by $(\mu, t, S)$ and $(\mu_o, t_o, S)$ are the same.

Let us conclude this section by considering again the examples.

Example 1 The extension of this exponential type family depends on the semigroup generated by the statistic $\{n, \Sigma_1^n t(x_i), \Sigma_1^n \ln h(x_i)\}$.

As an example consider the following special case of a normal distribution with positive mean and variance equal to 1.

$$f(x,\theta) = \frac{1}{\sqrt{2\pi}} e^{\theta x} e^{-\frac{1}{2}x^2} e^{-\frac{1}{2}\theta^2}, \quad x \in R,\ \theta > 0$$

We want to extend this family and the space S of multiples of likelihood functions consists of functions of the form

$$c\, e^{a\theta + b\theta^2}, \ c > 0,\ a \in R,\ b < 0.$$

The semigroup operations are given by

$$(a,\ b,\ c) \circ (a',\ b',\ c') = (a + a',\ b + b',\ cc')$$

and

$$\lambda * (a,\ b,\ c) = (a,\ b,\ \lambda c)$$

Homomorphisms of S have the form

$$h(a,\ b,\ c) = c^{\gamma}\ e^{a\alpha + b\beta}$$

$\alpha \in R$, $\beta \in R$, $\gamma \in R$. The homogeneous homomorphisms satisfy

$$h(a,\ b,\ \lambda c) = \lambda\ h(a,\ b,\ c)$$

and hence we have $\gamma = 1$.

For an element $h \in S^*$ we then have

$$h(t(x)) = h(x,\ -\tfrac{1}{2},\ \sqrt{2\pi}\ e^{-\frac{1}{2}x^2}) = \frac{1}{\sqrt{2\pi}}\ e^{-\frac{1}{2}x^2}\ e^{\alpha x - \frac{1}{2}\beta}\ .$$

When normalized this becomes the normal density with mean $\alpha \in R$. Thus the canonical family generated by the family with positive mean is the family with arbitrary mean.

Example 2 The space of likelihoods and their multiples are of the form

$$c\ 1_{[0,\theta]}(a), \quad c > 0,\ a > 0$$

with

$$(a,c) \circ (a',c') = (a \vee a',cc')$$

and

$$\lambda * (a,c) = (a,\lambda c)$$

A homomorphism has the form

$$h(a,c) = 1_{[0,\theta]}(a)\ c^{\alpha}$$

and the homogeneous ones have $\alpha = 1$. When normalized this gives the family we started with. Thus the uniform distributions on $[0,\theta]$, $0 < \theta < \infty$ form a canonical family.

Example 3 In this example the space S is spanned by the vectors

$$\left(\lambda \prod_{i=1}^{n} f_1(x_i),\ \lambda \prod_{i=1}^{n} f_2(x_i)\right), \quad \lambda > 0,\ x_i \in X,\ i = 1, \ldots ,n.$$

If we assume that any point is $]0,\infty[^2$ can be so represented then $S = ]0,\infty[^2$ with the operations

$$(a,b) \circ (a',b') = (aa',bb')$$

$$\lambda * (a,b) = (\lambda a,\lambda b)$$

A homomorphism is of the form

$$h(a,b) = a^{\alpha}\ b^{\beta}$$

and the homogeneous ones satisfy $\alpha + \beta = 1$.

Thus the family generated by $(f_1,f_2)$ has densities of the form

$$\frac{f_1^{\alpha}(x)\ f_2^{1-\alpha}(x)}{\int f_1^{\alpha}(x)\ f_2^{1-\alpha}(x)\ \mu(dx)}, \quad \alpha \in D,$$

$$D = \{\alpha \mid \int f_1^{\alpha}(x)\ f_2^{1-\alpha}(x)\ \mu(dx) < \infty\}.$$

## 5 References

1 Birkes, D., Brunk, H.D. and Lee, J.W.: Generalized Exponential Families. Tech.Rep.45, 1975, Oregon State University.

2 Johansen, S.: Exponential Models. Notes (1970). University of Copenhagen.

3 Lauritzen, S.L.: General Exponential Models for Discrete Observations, Scand.J.Statist.2: 23-33, 1975.

4 Loève, M.: Probability Theory. Van Nostrand, New York (1960).

PREPRINTS 1976

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Becker, Niels: Estimation for an Epidemic Model.

No. 2 Kaplan, Norman: A Generalization of a Result of Erdös and Renyi and a Related Problem.

No. 3 Lauritzen, Steffen & Remmer, Ole: An Investigation of the Asymptotic Properties of Geodetic Least Squares Estimates in Certain Types of Network.

No. 4 Asmussen, Søren: Almost Sure Behavior of Linear Functionals of Supercritical Branching Processes.

No. 5 Hald, Anders & Møller, Uffe: On the SPRT of the Mean of a Poisson Process.

No. 6 Lauritzen, Steffen: Projective Statistical Fields.

No. 7 Keiding, Niels: Inference and Tests for Fit in the Birth-Death-Immigration Process.

No. 8 Møller, Uffe: OC and ASN og the SPRT for the Poisson Process.

No. 9 Keiding, Niels: Population Growth and Branching Processes in Random Environments.

No. 10 Andersen, Per K., Andersen, Søren & Lauritzen, Steffen: The Average Noise from a Poisson Stream of Vehicles.

No. 11 Hald, Anders & Møller, Uffe: Multiple Sampling Plans of Given Strength for the Poisson and Binomial Distributions.

No. 12 Johansen, Søren: Two Notes on Conditioning and Ancillarity.

No. 13 Asmussen, Søren: Some Martingale Methods in the Limit Theory of Supercritical Branching Processes.

No. 14 Johansen, Søren: Homomorphisms and general exponential families.