CONVERGENCE ON MANIFOLDS OF GIBBS MEASURES WHICH ARE ABSOLUTELY CONTINUOUS WITH RESPECT TO HAUSDORFF MEASURES October 29, 2005

DENNIS D. COX, ROBERT HARDT, AND PETR KLOUČEK

ABSTRACT. We identify a sufficiant condition for a sequence of Gibbs measures P_λ with the density $Z_\lambda e^{-J_0(v)-J_1(v)}$, $v \in \mathbb{R}^n$, defined on a state space of $v's$, to converge weakly in a sense of measures to a Gibbs measure with a density $Z_\lambda e^{-J_0(v)}$, where the dominating measure for the density is the Hausdorff measure with an appropriate dimension. The function J_0 identifies an objective and J_1 defines a constraint. The condition we introduce requires the Hessian of J_1 to be non-negative definite and to have a constant rank on each component of $\{v \in \mathbb{R}^n \mid J_1(v) = 0\}$. The result presented shows that the probability measures P_λ concentrate on the highest dimensional stratum of $J_1^{-1}(0)$. We apply this result to a non-quasiconvex variational problem describing microstructural equilibria of a binary martensitic alloy. We show that the Young's measure describing, in general, non-attainable infimum of such a problem can be obtained as a "push-forward" measure induced by the probablity measure P_λ through a linear bounded operator $T_\lambda : GM \mapsto AY$, where GM denotes the space of Gibbs measures, and YM denotes the space of Young's measures defined as all probability measures generated by gradients of bounded sequences in a suitable Sobolev space.

1. INTRODUCTION

We consider some measure theoretic problems within the context of simulated annealing with constraints. Simulated annealing is a stochastic optimization algorithm that mimics the physical process of a system settling into the state of minimal energy. It is usually considered in a discrete state space setting when the objective has multiple optima, but continuous state space simulated annealing has found many applications ([2], [3], [7]). In a problem described below, we consider a discrete state space simulated annealing with, possibly, complicated constraints. The constraints can be expressed as the zero set of a nonnegative function, and then we can implement the algorithm through a relaxation method wherein we add in a nonnegative multiple of the constraint function to the objective and let the multiplier go to infinity. The results presented in this paper show that one must choose the constraint function carefully for otherwise the relaxation will introduce spurious terms into the objective.

The basic idea of Simulated Annealing (SA) goes back to [8], although it was not given its name until 30 years later [10]. Numerous advances in the general area of "Markov Chain Monte Carlo" (MCMC; see [12]) have lead to extensions of the basic SA algorithm which greatly improve its performance and range of applicability. Suppose our objective is to find

$$\arg\min_v J_0(v), \quad v \in \mathbb{R}^n.$$

Assume that J_0 is a sufficiently regular function (e.g., continuous), bounded below, and $J_0 \to \infty$ as $\|v\| \to \infty$ sufficiently fast that

$$\mathbb{Z}^{-1} = \int_{\mathbb{R}^n} e^{-J_0(v)}\, dv < \infty.$$

Key words and phrases. Gibbs Measure, Hausdorff Measure, Young's Measure, Crystalline microstructures.

Address: D. D. Cox, Department of Statistics, Rice University, 6100 Main Street, Houston, TX 77005, USA, E-mail: dcox@stat.rice.edu
Address: R. Hardt, Department of Mathematics, Rice University, 6100 Main Street, Houston, TX 77005, USA, E-mail: hardt@rice.edu
Address: P. Klouček, Department of Computational and Applied Mathematics, Rice University, 6100 Main Street, Houston, TX 77005, USA E-mail: kloucek@rice.edu.

Then, it is easy to construct Markov Chains so that the limiting (stationary) distribution of the chain will be the Borel probability measure with Lebesgue density function

$$f(v) \ = \ Ze^{-J_0(v)}. \tag{1.1}$$

Here, the limiting probability distribution is known as a *Gibbs measure*. The MCMC methodology provides many different approaches to obtain such chains with the Gibbs measure as limiting distribution. Now suppose that we wish to constrain v to the set

$$M \ = \ \{v : J_1(v) \ = \ 0\}\,, \tag{1.2}$$

where $J_1 \geq 0$ and also satisfies regularity conditions similarly to J_0. A convenient approach to solving the constrained minimization problem is to apply SA to the Gibbs measure with density

$$f(v;\lambda) \stackrel{\text{def}}{=} \ Z_\lambda e^{-J_0(v)-\lambda J_1(v)}, \tag{1.3}$$

and then let $\lambda \to \infty$. Here, Z_λ is a normalizing constant so that $\int f(v;\lambda)dv = 1$. We conjecture that the resulting limit would be the Gibbs measure with density (1.1) on M, where the dominating measure for the density would be Hausdorff measure on M of an appropriate dimension. The main result of this paper shows that, with a suitable condition on the second derivative of J_1, this conjecture is true.

2. Statement of the Main Result

We use *weak convergence of probability measures*. To define weak convergence of probability measures, suppose $\langle P_t : t \in \mathbb{Z}\rangle$ is a sequence of probability measures and P is a fixed probability measure, all defined on the Borel sets $\mathcal{E}$ of a given Polish space E. *Weak convergence of* P_n *to* P, denoted $P_n \Rightarrow P$, means

$$\int \phi dP_n \ \to \ \int \phi dP,$$

for all bounded continuous real valued functions ϕ. The theory of weak convergence of measures is presented in [1].

For $\lambda > 0$ and nonnegative continuous functions $J_0,\ J_1$ on $\mathbb{R}^n$ with

$$Z_\lambda^{-1} \ = \ \int_{\mathbb{R}^n} e^{-(J_0(x)+\lambda J_1(x))}\, dx \ < \ \infty\ ,$$

we have the probability density function

$$f(x;\lambda) \ = \ Z_\lambda e^{-(J_0(x)+\lambda J_1(x))}$$

and corresponding probability measure

$$P_\lambda(B) \ = \ \int_B f(x;\lambda)\, dx\ ,$$

defined for Lebesgue measurable subsets B of $\mathbb{R}^n$. We will investigate the behavior of P_λ as $\lambda \to \infty$.

Theorem 2.1. *Assume the following:*

(A1) $J_0 \in C^0(\mathbb{R}^n)$ *and* $J_1 \in C^3(\mathbb{R}^n)$.
(A2) $J_0 \geq 0$ *and* $J_1 \geq 0$.
(A3) *For some* $p > 0$, $J_0(x) \geq \|x\|^p$ *and* $J_1(x) \geq \|x\|^p$ *for all* $\|x\|$ *sufficiently large,* $p > 2$.
(A4) $M \ = \ \{x \in \mathbb{R}^n : J_1(x) = 0\}$ *is nonempty and bounded.*

(A5) *There exist bounded disjoint open set $U_1, U_2, \cdots, U_j$ and integers $0 \le k_1 < k_2 < \cdots < k_j \le n$ satisfying:*
   (a) $M \subset U = U_1 \cup U_2 \cup \cdots \cup U_j$,
   (b) *On each U_i, the Hessian $D^2 J_1$ is nonnegative definite and has constant rank k_i for $i = 1, \cdots, j$,*
   (c) *For some positive numbers C and α, $D^3 J_1$ satisfies the uniform Hölder condition*

$$\|D^3 J_1(x) - D^3 J_1(y)\| \le C\,|x-y|^\alpha \quad \textit{for} \quad x,\, y \in U\,.$$

Then each set $M_i = M \cap U_i$ is a $C^{2,\alpha}$ smooth $n - k_i$ dimensional manifold, and, as $\lambda \to \infty$, the probability measures converge

$$P_\lambda \Rightarrow P\,,$$

where, for any Borel set $B \subset \mathbb{R}^n$,

$$P(B) = Z \int_{M_1 \cap B} e^{-J_0(a)}\, \Lambda(a)^{-1/2}\, d\mathcal{H}^{n-k_1} a \quad \textit{with} \quad Z^{-1} = \int_{M_1} e^{-J_0}\, \Lambda^{-1/2}\, d\mathcal{H}^{n-k_1}\,, \tag{2.1}$$

$\Lambda(a)$ is the product of the k_1 positive eigenvalues of $D^2 J_1(a)$, and $\mathcal{H}^{n-k_1}$ is $(n-k_1)$-dimensional Hausdorff measure.

We note that the probability measures P_λ concentrate only on the highest dimensional stratum M_1 of $M = J_1^{-1}\{0\}$ and do not produce any lower dimensional measures on $M_2 \cup \cdots \cup M_j$.

In Section 7, we discuss a generalization of the main result, first to the case where the ambient space $\mathbb{R}^n$ is replaced by a compact Riemannian manifold and second to treat multiple limits

$$\lim_{\lambda_1 \to \infty} \lim_{\lambda_2 \to \infty} \cdots \lim_{\lambda_j \to \infty} P_{\lambda_1 \lambda_2 \cdots \lambda_j}$$

of probability measures in the form

$$P_{\lambda_1 \lambda_2 \cdots \lambda_j} = Z_{\lambda_1 \lambda_2 \cdots \lambda_j} \int_{\mathbb{R}^n} e^{-W_0(x) - \lambda_1(x) W_1(x) - \lambda_2(x) W_2(x) - \cdots - \lambda_j(x) W_j(x)}\, dx$$

Sections 3 through 5 gather material necessary to prove Theorem 2.1 in Section 6.

3. Nearest Point Projection for a Submanifold

Recall that, for any Borel set $A \subset \mathbb{R}^n$ and $0 \le k \le n$, the k dimensional Hausdorff measure $\mathcal{H}^k(A)$ is defined [6], 2.10.2. It is normalized so that, for integer k, in $\mathbb{R}^k$, $\mathcal{H}^k$ coincides with k dimensional Lebesgue measure. In a higher dimensional $\mathbb{R}^n$, the restriction of $\mathcal{H}^k$ to a k dimensional C^1 submanifold M coincides with the Riemannian volume measure on M for the metric induced from $\mathbb{R}^n$. In particular, a k dimensional ball of radius r in $\mathbb{R}^k$,

$$\mathbb{B}^k_r(a) \equiv \{x \in \mathbb{R}^k : |x-a| < r\}\,,$$

has

$$\mathcal{H}^k(\mathbb{B}^k_r(a)) = \alpha_k r^k \tag{3.1}$$

where α_k is the k dimensional Lebesgue measure of the unit ball in $\mathbb{R}^k$.

Our notation for an integral with respect to a (lower dimensional) Hausdorff measure will have the form

$$\int_A f(a)\, d\mathcal{H}^k a \quad \text{or} \quad \int_A f\, d\mathcal{H}^k\,,$$

while our integrals with respect to the (top dimensional) Lebesgue will keep the standard notation

$$\int_U f(x)\, dx \qquad \text{rather than} \qquad \int_U f(x)\, d\mathcal{H}^n x.$$

In particular, we have the polar coordinate formula for a Lebesgue integrable function f on the ball $\mathbb{B}_R(0) \equiv \mathbb{B}^n_R(0)$

$$\int_{\mathbb{B}_R(0)} f(x)\, dx = \int_{S^{n-1}} \int_0^R f(r\omega)\, r^{n-1}\, dr\, d\mathcal{H}^{n-1}\omega\,,$$

where S^{n-1} is the $n-1$ dimensional unit sphere in $\mathbb{R}^n$. One readily checks that

$$\mathcal{H}^{n-1}(S^{n-1}) \;=\; n\,\alpha_n \tag{3.2}$$

by differentiating (3.1).

For any vector subspace T of $\mathbb{R}^n$, the orthogonal projection

$$\Pi_T \;:\; \mathbb{R}^n \;\to\; T$$

is the linear map which takes any point $x \in \mathbb{R}^n$ to the unique point $\Pi_T(x)$ in T that is *nearest* to x.

Suppose that M is a compact m dimensional $C^{2,\alpha}$ submanifold of $\mathbb{R}^n$. Then the

$$m \text{ dimensional } \mathit{tangent\ space}\;\; T_aM \quad \text{and} \quad n-m \text{ dimensional } \mathit{normal\ space}\;\; (T_aM)^{\perp}$$

are continuously differentiable functions of $a \in M$. For the submanifold M there is also a nearest point map Π_M that is well-defined in some "tubular" neighborhood of M. Its differential at a point x is close to the orthogonal projection of $\mathbb{R}^n$ onto $T_{\Pi_M(x)}M$. Specifically, we have the

Lemma 3.1. *There is a bounded open neighborhood U of M in $\mathbb{R}^n$ so that every point $x \in U$ has a unique nearest point $\Pi_M(x)$ in M. Moreover, on some such U, the map Π_M is $C^{1,\alpha}$ smooth, and there is a positive constant C so that*

$$\|D\Pi_M(x) \;-\; \Pi_{T_{\Pi_M(x)}M}\| \;\le\; C|\Pi_M(x) - x|^{\alpha}$$

for all $x \in U$.

Proof. As discussed for example in [5], the nearest point neighborhood property of a compact C^2 submanifold M depends on its curvature bound. It holds specifically for the open neighborhood

$$\{x \in \mathbb{R}^n \;:\; \text{dist}(x, M) < (\max_{a\in M} \|A_M(a)\|)^{-1}\}\ ,$$

where A_M is the second fundamental form of M. For $a \in M$, $D\Pi_M(a) = \Pi_{T_aM}$. Since, for a compact $C^{2,\alpha}$ submanifold M, the map Π_M is $C^{1,\alpha}$ bounded in some compact neighborhood of M, the desired estimate of $\|D\Pi_M(x) - \Pi_{T_{\Pi_M(x)}M}\|$ follows. ☐

Corollary 3.2. *The m dimensional Jacobian $J_m\Pi_M \equiv \|\wedge_m D\Pi_M\| = \sqrt{\det\left((D\Pi_M)\circ(D\Pi_M)^*\right)}$ (See [6] 3.2.22) satisfies*

$$|\; J_m\Pi_M(x) \;-\; 1\;| \;\le\; C|\Pi_M(x) - x|^{\alpha}$$

on some such neighborhood U for some positive constant C.

Proof. The linear map $\Pi_{T_{\Pi(x)}M}$, being an orthogonal projection onto an m dimensional space, has m Jacobian 1. Since $\sqrt{t}$ is smooth near $t = 1$, the estimate follows from the formula for J_m and Lemma 3.1. ☐

For small U, each slice $\Pi_M^{-1}\{a\}$ is the graph of a C^1 small function over an $n-m$ dimensional normal disk

$$N_\delta(a) \;\equiv\; \mathbb{B}_\delta(0) \cap (T_aM)^{\perp}\ .$$

More precisely,

Corollary 3.3. *For some such U, there are positive δ and C and, for every point $a \in M$, a $C^{1,\alpha}$ function*

$$g_a \;:\; N_\delta(a) \;\longrightarrow\; T_aM$$

so that $g_a(0) = 0$, $\|Dg_a(y)\| \le C|y|^{\alpha}$ for $y \in N_\delta(a)$, and

$$\Pi_M^{-1}\{a\} \;=\; G_a(N_\delta(a)) \quad \textit{where} \quad G_a(y) \;=\; a + y + g_a(y)\ .$$

It follows that $G_a : N_\delta(a) \to \Pi_M^{-1}\{a\}$ is a C^1 diffeomorphism satisfying

$$| \, J_{n-m} G_a(y) \; - \; 1 \, | \; \leq \; C|y|^{\alpha} \; ,$$

where $J_{n-m}G_a \equiv \| \wedge_{n-m} DG_a \| = \sqrt{\det\left((DG_a)^* \circ (DG_a)\right)}$.

For $x \in U$ *and* $a = \Pi_M(x)$*, the inverse relation*

$$x \; = \; G_a(y) \quad \Leftrightarrow \quad y \; = \; \Pi_{(T_aM)^{\perp}}(x-a)$$

shows that

$$|y| \; \leq \; |x-a| \; \leq \; |y| \, + \, C|y|^{1+\alpha} \; . \tag{3.3}$$

Proof. By Lemma 3.1, the rank of $D\Pi_M(x)$ equals m for x near M. By the rank theorem [4], 10.3, the set $\Pi_M^{-1}\{a\}$ is a $C^{1,\alpha}$ submanifold orthogonal to M at a, and hence the graph of a $C^{1,\alpha}$ function over a ball in the normal space to M at a. As a varies over the compact submanifold M, estimates on the size of this ball and the $C^{1,\alpha}$ norm of this function are all uniform, by the proof of the rank theorem. □

Integrals over such a tubular neighborhood may be computed using the above Jacobians:

Lemma 3.4. *For any bounded continuous function* ψ *on* $\mathbb{R}^n$,

$$\int_U \psi(x)\,dx \; = \; \int_M \left(\int_{\Pi_M^{-1}\{a\}} \psi \cdot (J_m\Pi_M)^{-1}\, d\mathcal{H}^{n-m} \right) d\mathcal{H}^m a$$

$$= \; \int_M \left(\int_{N_\delta(a)} \psi(G_a(y)) \cdot (J_m\Pi_M)^{-1}(G_a(y)) \cdot J_{n-m}(G_a(y))\,dy \right) d\mathcal{H}^m a \, .$$

Proof. For the first equality, we may apply the (coarea) change of variable formula [6], 3.2.22(3) for the map $\Pi_M : U \to M$,

$$\int_U \phi(x)(J_m\Pi_M)(x)\,dx \; = \; \int_M \left(\int_{\Pi_M^{-1}\{a\}} \phi\, d\mathcal{H}^{n-m} \right) d\mathcal{H}^m a \; ,$$

with $\phi(x) = \psi(x) \cdot (J_m\Pi_M)^{-1}(x)$. For the second equality, we then apply the (area) change of variable formula [6], 3.2.5 for the map $G_a : N_\delta(a) \to \Pi_M^{-1}\{a\}$,

$$\int_{N_\delta(a)} (\phi(G_a(y)) \cdot (J_{n-m}G_a)(y)\,dy \; = \; \int_{\Pi_M^{-1}\{a\}} \phi\, d\mathcal{H}^{n-m}$$

with $\phi = \psi \cdot (J_m\Pi_M)^{-1}$. □

For use in Section 7, we next observe that:

All the results through Corollary 3.2 continue to hold in case the ambient space $\mathbb{R}^n$ *is replaced by an* n *dimensional Riemannian manifold* N.

Concerning Lemma 3.4, one additional observation is required. In the $\mathbb{R}^n$ case, each set

$$\Pi_M^{-1}\{a\} \; = \; \{x \in \mathbb{R}^n \; : \; \mathrm{dist}(x,M) < \delta, \; \Pi_M(x) = a\} \; = \; \{a + v \; : \; v \in (T_M(a))^{\perp}, \; |v| < \delta\} \tag{3.4}$$

is simply a flat $n-m$ dimensional planar disk, while, in the general Riemannian case, it is a uniformly smooth (but possibly curved) $n-m$ dimensional disk in N.

More precisely, consider, for each $a \in M$, the $n-m$ dimensional planar disk in the normal space

$$V_\delta(a) \; \equiv \; \{v \in (T_aM)^{\perp} \; : \; |v| < \delta\} \; \subset \; T_aN \; .$$

Lemma 3.5. *There exist positive δ and C and, for every point $a \in M$, a $C^{1,\alpha}$ function G_a mapping $V_\delta(a)$ diffeomorphically onto $\Pi_M^{-1}\{a\}$ so that $G_a(0) = a$, $DG_a(0)$ is an isometry, and, for every $y \in V_\delta(a)$,*

$$\|DG_a(y) - DG_a(0)\| \;\le\; C|y|^\alpha;$$

hence,

$$|y| \;\le\; \mathit{dist}(G_a(y), a) \;\le\; |y| \,+\, C|y|^{1+\alpha}\ , \tag{3.5}$$

and

$$|\, J_{n-m}G_a(y) \;-\; 1 \,| \;\le\; C|y|^\alpha\ ,$$

where $J_{n-m}G_a = \sqrt{\det\left((DG_a)^* \circ (DG_a)\right)}$.

Proof. In the general Riemannian case, the set $\Pi_M^{-1}\{a\}$ is a totally geodesic $n-m$ dimensional disk in N that is orthogonal to M at a. The desired parameterizing map G_a is obtained by simply restricting the Riemannian exponential map Exp_a^N to the normal disk $V_\delta(a)$. The estimates all follow from properties of this exponential map. In particular, as a varies over the compact submanifold M, all estimates are uniform because of the $C^{0,\alpha}$ bound on the sectional curvature of M. □

Lemma 3.4 is now replaced by:

Lemma 3.6. *For any bounded continuous function ψ on N,*

$$\int_U \psi(x)\,dx \;=\; \int_M \left(\int_{\Pi_M^{-1}\{a\}} \psi\cdot(J_m\Pi_M)^{-1}\,d\mathcal{H}^{n-m}\right) d\mathcal{H}^m(a)$$

$$=\; \int_M \left(\int_{V_\delta(a)} \psi(G_a(y))\cdot(J_m\Pi_M)^{-1}(G_a(y))\cdot J_{n-m}(G_a(y))\,dy\right) d\mathcal{H}^m(a)\,.$$

Proof. The first equality follows from [6], 3.2.22(3) as in the proof of 3.4. For the second equality, we then apply the (area) change of variable formula [6], 3.2.5 for the map $G_a : V_\delta(a) \to \Pi_M^{-1}\{a\}$,

$$\int_{V_\delta(a)} \phi(G_a(y)\cdot(J_{n-m}G_a)(y)\,dy \;=\; \int_{\Pi_M^{-1}\{a\}} \phi\,d\mathcal{H}^{n-m}$$

with $\phi = \psi\cdot(J_m\Pi_M)^{-1}$. □

4. The Zero Set of a Nonnegative Function of Fixed Nondegeneracy

Theorem 4.1. *Suppose that F is a nonnegative, $C^{3,\alpha}$ smooth function on $\mathbb{R}^n$, $m \in \{0,\cdots,n-1\}$, $M = F^{-1}\{0\}$ is compact, and rank$\frac{\partial^2 F}{\partial x_i \partial x_j}(x) \equiv n-m$ for all x in some neighborhood of M. Then, there are positive numbers δ and C so that:*

(1) For $a \in M$, grad$F(a)$ vanishes, and the symmetric matrix $\frac{\partial^2 F}{\partial x_i \partial x_j}(a)$ has, counting multiplicities, m zero eigenvalues and $n-m$ positive eigenvalues

$$\lambda_1(a) \;\le\; \lambda_2(a) \;\le\; \cdots \;\le\; \lambda_{n-m}(a)\ ,$$

which are continuous in a with a positive minimum and a finite maximum.

(2) M is an m dimensional embedded $C^{2,\alpha}$ smooth submanifold.

(3) For each $a \in M$, there is an an orthogonal rotation Γ_a of $\mathbb{R}^n$ so that

$$\Gamma_a(\{0\}\times\mathbb{R}^{n-m}) \;=\; (T_aM)^\perp$$

and

$$|F(a+\Gamma_a(x)) \;-\; \sum_{i=m+1}^{n} \frac{1}{2}\lambda_i(a)(x_i-a_i)^2| \;\le\; C(\sup_{\mathbb{B}_\delta(a)} \|D^2F\|)|x-a|^3$$

for all $x \in \mathbb{B}_\delta(a)$.

Note that the numbers δ and C in Theorem 4.1 as well as the δ and C of Corollary 3.3 are all uniform, *independent of* $a \in M$.

Proof of 4.1. To verify (1), note that for each $a \in M$ and each vector $v \in \mathbb{R}^n$, the function $F_v(t) = F(a+tv)$ has a minimum at $t=0$. So $v{\cdot}\mathrm{grad}F(a) = \frac{dF_v}{dt}|_{t=0} = 0$, and

$$0 \;\le\; \frac{d^2F_v}{dt^2}|_{t=0} \;=\; \frac{d}{dt}|_{t=0} v\cdot \mathrm{grad}F(a+tv) \;=\; \sum_{i,j} v_i \frac{\partial^2 F}{\partial x_i \partial x_j}(a) v_j \;.$$

Thus $\mathrm{grad}F(a) = 0$, and all the eigenvalues of $\frac{\partial^2 F}{\partial x_i \partial x_j}(a)$ are nonnegative. In general, the full collection of eigenvalues of a square matrix, being the complex roots of the characteristic polynomial, varies continuously as the matrix varies. See e.g. [6] 4.3.??. Here, by assumption, for $a \in M$, the matrix $\frac{\partial^2 F}{\partial x_i \partial x_j}(a)$ has, counting multiplicities, precisely m zero eigenvalues and precisely $n-m$, nonzero, hence positive, eigenvalues. So, under ordering by size, these positive eigenvalues become continuous functions on M. By the compactness of M, λ_1 has a positive minimum and λ_{n-m} a finite maximum.

To verify (2), we first observe that, near each point $a \in M$, the set $\tilde{M} = (\mathrm{grad}F)^{-1}\{(0,\cdots,0)\}$ is a $C^{2,\alpha}$ smooth submanifold of dimension m. This follows from the general rank theorem [4], 10.3. So it suffices to show that, locally near any such a, M coincides with $\tilde{M}$. To see this, consider a connected coordinate neighborhood Ω of a in $\tilde{M}$. For $b \in \tilde{V}$, there is a C^1 curve $\gamma : [0,1] \to \Omega$. Since

$$F(b) \;=\; F(b)-F(a) \;=\; \int_0^1 (F\circ\gamma)'(t)\,dt \;=\; \int_0^1 (\mathrm{grad}F)(\gamma(t))\cdot\gamma'(t)\,dt \;=\; \int_0^1 0\,dt \;=\; 0\,,$$

b belongs to M. Thus $\Omega \subset M$. We conclude that $M \cap \tilde{M}$ is open, as well as closed, relative to $\tilde{M}$ so that, near M, M coincides with $\tilde{M}$.

For (3), we let $v_1,\cdots,v_n$ be orthonormal eigenvectors of the symmetric matrix $\frac{\partial^2 F}{\partial x_i \partial x_j}(a)$ corresponding to the eigenvalues $0,\cdots,0,\lambda_1(a),\cdots,\lambda_{n-m}(a)$, and choose the rotation Γ_a of $\mathbb{R}^n$ satisfying $\Gamma_a(\mathbf{e}_i) = v_i$ for $i=1,\cdots,n$. With $H_a(x) = a+\Gamma_a(x)$, we deduce that

$$\frac{\partial^2(F\circ H_a)}{\partial x_i \partial x_j}(0)[\mathbf{e}_i] \;=\; \begin{cases} 0 & \text{for } \; i=1,\cdots,m \\ \lambda_i(a)\mathbf{e}_i & \text{for } \; i=m+1,\cdots,n\,. \end{cases}$$

Since also $(F\circ H_a)(0) = 0$, $\mathrm{grad}(F\circ H_a)(0) = 0$, and $\|D^2F\|$ is bounded in some neighborhood of M, the second order Taylor expansion for $F\circ G_a$ now gives (3). □

5. Some Integrals.

Lemma 5.1. *For* $k = 1,2,\cdots$,

$$\int_0^\infty e^{-\lambda t^2} t^{k-1}\,dt \;=\; \beta_k \lambda^{-\frac{k}{2}}$$

where

$$\beta_k \;=\; \begin{cases} 2^{-\frac{k}{2}}(k-2)(k-4)\cdots(2) & \textit{for } k \textit{ even} \\ 2^{-\frac{k-1}{2}}(k-2)(k-4)\cdots(3)\cdot\sqrt{\pi} & \textit{for } k \textit{ odd}\,. \end{cases}$$

Proof. The substitution $s = \sqrt{\lambda}t$ gives the factor $\lambda^{-\frac{k}{2}}$ and reduces to the case $\lambda = 1$.

Integration by parts gives

$$\int_0^\infty e^{-t^2} t^{k-1}\,dt \;=\; \frac{-1}{2}\int_0^\infty t^{k-2}\,d(e^{-t^2}) \;=\; \frac{k-2}{2}\int_0^\infty e^{-t^2}t^{k-3}\,dt\,.$$

This may be applied with k replaced by $k-2, k-4, \cdots$, finally giving the formula

$$\int_0^\infty e^{-t^2}t^{2j}\,dt \;=\; \begin{cases} 2^{-\frac{k-2}{2}}(k-2)(k-4)\cdots(2)\cdot\int_0^\infty e^{-t^2}t\,dt & \text{for } k \text{ even}\\ 2^{-\frac{k-1}{2}}(k-2)(k-4)\cdots(3)\cdot\int_0^\infty e^{-t^2}\,dt & \text{for } k \text{ odd}\,.\end{cases}$$

Of course, substituting $s = t^2$ gives $\int_0^\infty e^{-t^2}t\,dt = \frac{1}{2}$, and the last integral is found by the usual polar coordinate trick

$$\left(\int_0^\infty e^{-t^2}\,dt\right)^2 \;=\; \int_0^\infty\int_0^\infty e^{-x^2-y^2}\,dx\,dy \;=\; \int_0^{2\pi}\int_0^\infty e^{-r^2}r\,dr\,d\theta \;=\; \frac{2\pi}{2}\int_0^\infty e^{-u}\,du \;=\; \pi\,.$$

□

Corollary 5.2.

$$\lim_{\lambda\to\infty}\lambda^{\frac{k}{2}}\int_0^\infty e^{-\lambda t^2}t^j\,dt \;=\; 0 \quad \textit{for any integer } j > k-1\,, \tag{5.1}$$

and

$$\lim_{\lambda\to\infty}\lambda^{\frac{k}{2}}\int_\delta^\infty e^{-\lambda t^2}t^{k-1}\,dt \;=\; 0 \quad \textit{for any } \delta > 0 \tag{5.2}$$

Proof. Applying Lemma 5.1 with $k = j+1$ gives the first conclusion because $\lambda^{(-j+k-1)/2} \to 0$ as $\lambda\to\infty$. For the second, we change variables $s = \lambda^{1/2}t$ to see that

$$\lambda^{\frac{k}{2}}\int_\delta^\infty e^{-\lambda t^2}t^{k-1}\,dt \;=\; \lambda^{\frac{k}{2}}\lambda^{-\frac{k-1}{2}}\lambda^{-1/2}\int_{\lambda^{1/2}\delta}^\infty e^{-s^2}s^{k-1}\,ds \;\to\; 0 \text{ as } \lambda\to\infty$$

because $\int_0^\infty e^{-s^2}s^{k-1}\,ds < \infty$. □

Corollary 5.3. *For* $\delta > 0$ *and* $0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_k < \infty$*,*

$$\lim_{\lambda\to\infty}\lambda^{\frac{k}{2}}\int_{\mathbb{B}^k_\delta(0)} e^{-\frac{1}{2}\lambda(\lambda_1 y_1^2+\cdots+\lambda_k y_k^2)}\,dy \;=\; 2^{k/2}\Lambda^{-1/2}k\alpha_k\beta_k$$

where $\Lambda = \lambda_1\cdots\lambda_k$.

Proof. One can explicitly compute the integral over the k dimensional elliptical region

$$E^k_\delta \;=\; \{\, y\in\mathbb{R}^k \;:\; \lambda_1 y_1^2+\cdots+\lambda_1 y_k^2 \;<\; 2\delta^2\,\}\,,$$

because $E^k_\delta = L(\mathbb{B}^k_\delta(0))$ where

$$L(z_1,\cdots,z_k) \;=\; ((2/\lambda_1)^{1/2}z_1,\cdots,(2/\lambda_k)^{1/2}z_k) \quad \text{for } (z_1,\cdots,z_k)\in\mathbb{R}^k\,.$$

Using the change of variables $y = L(z)$ with $dy = (J_kL)\,dz = 2^{k/2}\Lambda^{-1/2}dz$, as well as polar coordinates, Lemma 5.1, and equation (5.2), we find that

$$
\begin{aligned}
\lambda^{\frac{k}{2}} \int_{E^k_\delta} e^{-\frac{1}{2}\lambda(\lambda_1 y_1^2+\cdots+\lambda_k y_k^2)}\, dy \;&=\; 2^{k/2}\Lambda^{-1/2}\lambda^{\frac{k}{2}} \int_{B^k_\delta} e^{-\lambda|z|^2}\, dz \\
&=\; 2^{k/2}\Lambda^{-1/2}\lambda^{\frac{k}{2}} \int_{S^{k-1}}\int_0^\delta e^{-\lambda r^2} r^{k-1}\, dr\, d\mathcal{H}^{k-1} \\
&=\; 2^{k/2}\Lambda^{-1/2}\lambda^{\frac{k}{2}} k\alpha_k \int_0^\delta e^{-\lambda r^2} r^{k-1}\, dr \\
&=\; 2^{k/2}\Lambda^{-1/2} k\alpha_k \lambda^{\frac{k}{2}} [\int_0^\infty e^{-\lambda r^2} r^{k-1}\, dr \;-\; \int_\delta^\infty e^{-\lambda r^2} r^{k-1}\, dr\,] \\
&\to 2^{k/2}\Lambda^{-1/2} k\alpha_k\beta_k \;-\; 0 \qquad \text{as } \lambda\to\infty\,.
\end{aligned}
$$

We get precisely the same limit with E^k_δ replaced by the ball $\mathbb{B}^k_\delta(0)$ because we have the inclusions

$$
\mathbb{B}^k_\gamma(0) \subset \mathbb{B}^k_\delta(0) \subset \mathbb{B}^k_\epsilon(0) \qquad \text{and} \qquad \mathbb{B}^k_\gamma(0) \subset E^k_\delta(0) \subset \mathbb{B}^k_\epsilon(0)
$$

with $\gamma = \min\{\delta, (2/\lambda_k)^{1/2}\delta\}$ and $\epsilon = \max\{\delta, (2/\lambda_1)^{1/2}\delta\}$ and we have

$$
\begin{aligned}
\lim_{\lambda\to\infty} \lambda^{\frac{k}{2}} \int_{\mathbb{B}^k_\epsilon(0)\setminus\mathbb{B}^k_\gamma(0)} e^{-\frac{1}{2}\lambda(\lambda_1 y_1^2+\cdots+\lambda_k y_k^2)}\, dy \;&\le\; \lim_{\lambda\to\infty} \lambda^{\frac{k}{2}} \int_{\mathbb{B}^k_\epsilon(0)\setminus\mathbb{B}^k_\gamma(0)} e^{-\frac{1}{2}\lambda(\lambda_1|y|^2)}\, dy \\
&\le \lim_{\lambda\to\infty} k\alpha_k\lambda^{\frac{k}{2}} \int_\gamma^\infty e^{-\frac{1}{2}\lambda(\lambda_1 r^2)} r^{k-1}\, dr = \; 0
\end{aligned}
$$

by equation (5.2). $\square$

Lemma 5.4. *Suppose $0 \le k \le n$, F is a nonnegative continuous function on $\mathbb{R}^n$, $p > 0$, and $F(y) \ge |y|^p$ whenever $|y|$ is sufficiently large. Then, for any bounded open neighborhood U of $F^{-1}\{0\}$,*

$$
\lim_{\lambda\to\infty} \lambda^{k/2} \int_{\mathbb{R}^n\setminus U} e^{-\lambda F(y)}\, dy \;=\; 0\,.
$$

Proof. We may assume $p < 2$. Choose $R > 0$ so that $\overline{U} \subset \mathbb{B}_R(0)$ and $F(y) \ge |y|^p$ whenever $|y| \ge R$. On the bounded region $\mathbb{B}_R(0)\setminus U$, F has a positive lower bound ε, and

$$
\lambda^{k/2} e^{-\lambda F(y)} \;\le\; \lambda^{k/2} e^{-\lambda\varepsilon} \;\to\; 0
$$

uniformly as $\lambda\to\infty$. Thus

$$
\lim_{\lambda\to\infty} \lambda^{k/2} \int_{\mathbb{B}_R(0)\setminus U} e^{-\lambda F(y)}\, dy \;=\; 0\,.
$$

For the remaining set $\mathbb{R}^k \setminus \mathbb{B}_R(0)$, we use polar coordinates and change variables $s = \lambda^{1/p} t$ to see that

$$
\lambda^{k/2} \int_{\mathbb{R}^n\setminus\mathbb{B}_R(0)} e^{-\lambda F(y)}\, dy \;\le\; n\alpha_{n-1}\lambda^{k/2} \int_R^\infty e^{-\lambda t^p} t^{n-1}\, dt \;=\; n\alpha_{n-1}\lambda^{\frac{k}{2}-\frac{n-1}{p}-\frac{1}{p}} \int_{\lambda^{1/p}R}^\infty e^{-s^p} s^{n-1}\, ds \;\to\; 0
$$

as $\lambda\to\infty$ because $\frac{k}{2}-\frac{n}{p} \le 0$ and $\int_0^\infty e^{-s^p} s^{n-1}\, ds < \infty$. $\square$

6. Proof of Main Theorem.

First we treat the case $j = 1$ where the Hessian $D^2 J_1$ has constant rank $k = k_1$ in an open set U_1 containing all of $M = J_1^{-1}\{0\}$ so that M is, by Theorem 4.1, an $n-k$ dimensional submanifold of $\mathbb{R}^n$.

Taking $F = J_1$ and $M = J_1^{-1}\{0\}$, we choose δ and $U \subset U_1$ small enough and C large enough to satisfy Lemma 3.1, Corollary 3.2, Corollary 3.3, and Theorem 4.1 with $m = n-k$. In the remainder of the proof, we will occasionally enlarge C, finitely many times, without changing the notation. Nevertheless, the constant C will always just depend on n and J_1.

Let ϕ be a bounded continuous function on $\mathbb{R}^n$ and ε be any positive number satisfying

$$\varepsilon \ < \ \min\left\{\frac{1}{2}, \frac{1}{2C}\right\} \ ,$$

with C as in Corollary 3.2.

First we may assume that the tubular neighborhood U in Lemma 3.1, Corollary 3.2, and Corollary 3.3 (with the corresponding δ from Corollary 3.3) is small enough so that, for any point $x \in U$ and nearest point $a = \Pi_M(x) \in M$, one has

$$|x - a| \ < \ \varepsilon \ , \quad |J_0(x) - J_0(a)| \ < \ \varepsilon \ , \quad |\phi(x) - \phi(a)| \ < \ \varepsilon \ . \tag{6.1}$$

For λ sufficiently large, we have, by Lemma 5.4, that

$$\lambda^{\frac{k}{2}} \int_{\mathbb{R}^n \setminus U} \phi(x) e^{-J_0(x) - \lambda J_1(x)} \, dx \ \le \ \lambda^{\frac{k}{2}} \int_{\mathbb{R}^n \setminus U} e^{-\lambda J_1(x)} \, dx \ < \ \varepsilon \ . \tag{6.2}$$

For the integral over U, we use Lemma 3.4 with ψ replaced by $\phi e^{-J_0 - \lambda J_1}$ to write

$$\begin{aligned} &\lambda^{\frac{k}{2}} \int_U \phi(x) e^{-J_0(x) - \lambda J_1(x)} \, dx \\ &= \lambda^{\frac{k}{2}} \int_M \left(\int_{N_{\delta}(a)} \phi(G_a(y)) \cdot e^{-J_0(G_a(y))} \cdot e^{-\lambda J_1(G_a(y))} \cdot (J_{n-k}\Pi_M)^{-1}(G_a(y)) \cdot J_k(G_a(y)) \, dy \right) d\mathcal{H}^{n-k} a \, . \end{aligned}$$

We may make upper estimates on each of these five terms in the integrand. Note that, by our choice of U and (3.3), with $x = G_a(y) \in U$ (hence $a = \Pi_M(x)$), one also has

$$|y| \ \le \ |x - a| \ < \ \varepsilon \ .$$

For the first term we use (6.1) to see that

$$\phi((G_a(y)) \ \le \ \phi(a) \ + \ \varepsilon \ .$$

For the second term we again use (6.1)

$$e^{-J_1((G_a(y))} \ \le \ e^{-J_1(a) + \varepsilon} \ = \ e^{-J_1(a)} e^{\varepsilon} \ \le \ e^{-J_1(a)}(1 + 2\varepsilon) \ .$$

For the fourth term, we use Corollary 3.2 to infer that

$$(J_{n-k}\Pi_M)^{-1}(G_a(y)) \ \le \ (1 - C\varepsilon^{\alpha})^{-1} \ \le \ 1 + 2C\varepsilon^{\alpha} \ .$$

For the fifth term, we use Corollary 3.3 to obtain

$$J_k(G_a(y)) \ \le \ (1 + C\varepsilon^{\alpha}) \ .$$

Combining these and changing C, we now have the upper bound

$$\lambda^{\frac{k}{2}} \int_U \phi(x) e^{-J_0(x) - \lambda J_1(x)} \, dx \le \lambda^{\frac{k}{2}} (1 + C\varepsilon^{\alpha}) \int_M (\phi(a) + \varepsilon) e^{-J_0(a)} \int_{N_{\delta}(a)} e^{-\lambda J_1(G_a(y))} \, dy \, d\mathcal{H}^{n-k} a \, . \tag{6.3}$$

The remaining third term is in the integral over $N_\delta(a)$. To estimate this, we rotate coordinates as in Theorem 4.1(3) with $F = J_1$ and use (6), Corollary 5.3 with $k = k$, and (5.1) with $k = k + 2$ to deduce that

$$\begin{aligned} \lambda^{\frac{k}{2}} \int_{N_\delta(a)} e^{-\lambda J_1(G_a(y))}\, dy \quad &\leq \quad \lambda^{\frac{k}{2}} \int_{N_\delta(a)} e^{-\lambda(\sum_{i=1}^k \frac{1}{2}\lambda_i(a) y_i^2)} e^{C\lambda|y|^3}\, dy \\ \leq\ \lambda^{\frac{k}{2}} \int_{N_\delta(a)} e^{-\lambda(\sum_{i=1}^k \frac{1}{2}\lambda_i(a) y_i^2)} (1\ +\ 2C\lambda|y|^3)\, dy \\ \leq\ \lambda^{\frac{k}{2}} \int_{N_\delta(a)} e^{-\lambda(\sum_{i=1}^k \frac{1}{2}\lambda_i(a) y_i^2)}\, dy\ +\ 2C\lambda^{\frac{k+2}{2}}\, k\alpha_k \int_0^\infty e^{-\frac{1}{2}\lambda\lambda_1(a) r^2} r^{3+k-1}\, dr \\ \longrightarrow\ 2^{\frac{k}{2}} \Lambda^{-1/2}(a)\, k\alpha_k\beta_k \quad + \quad 0 \quad \text{as } \lambda \to \infty\ , \end{aligned} \tag{6.4}$$

where $\Lambda(a) \ = \ \lambda_1(a) \cdot \cdots \cdot \lambda_k(a)$. Taking the $\limsup_{\lambda\to\infty}$ in (6.3) along with (6.4), recalling (6.2), and then letting $\varepsilon \downarrow 0$, we conclude that

$$\limsup_{\lambda\to\infty} \lambda^{\frac{k}{2}} \int_{\mathbb{R}^n} \phi(x) e^{-J_0(x)-\lambda J_1(x)}\, dx \ \leq\ 2^{\frac{k}{2}}\, k\alpha_k\beta_k \int_M e^{-J_0(a)} \Lambda^{-1/2}(a)\, d\mathcal{H}^{n-k} a\ . \tag{6.5}$$

Next, arguing in the same manner using lower bounds gives the inequality

$$\liminf_{\lambda\to\infty} \lambda^{\frac{k}{2}} \int_{\mathbb{R}^n} \phi(x) e^{-J_0(x)-\lambda J_1(x)}\, dx \ \geq\ 2^{\frac{k}{2}}\, k\alpha_k\beta_k \int_M e^{-J_0(a)} \Lambda^{-1/2}(a)\, d\mathcal{H}^{n-k} a\ . \tag{6.6}$$

This essentially finishes the proof. For the normalization, we define, for $\lambda > 1$,

$$Y_\lambda \ = \left(2^{\frac{k}{2}}\, k\alpha_k\beta_k\right)^{-1} Z\lambda^{\frac{k}{2}}\ ,$$

where

$$Z \ = \left(\int_M e^{-J_0(a)} \Lambda^{-1/2}(a)\, d\mathcal{H}^{n-k} a\right)^{-1}$$

as before. We now apply (6.5) and (6.6) first with $g \equiv 1$ to see that

$$\lim_{\lambda\to\infty} \frac{Y_\lambda}{Z_\lambda} \ = \ \lim_{\lambda\to\infty} Y_\lambda \int_{\mathbb{R}^n} e^{-J_0(x)-\lambda J_1(x)}\, dx \ = \ 1\ ,$$

and second with the general bounded continuous ϕ to obtain the conclusion

$$\begin{aligned} &\lim_{\lambda\to\infty} Z_\lambda \int_{\mathbb{R}^n} \phi(x) e^{-J_0(x)-\lambda J_1(x)}\, dx \\ = \ \lim_{\lambda\to\infty} Y_\lambda \int_{\mathbb{R}^n} \phi(x) e^{-J_0(x)-\lambda J_1(x)}\, dx \ &= \ Z \int_M \phi(a) e^{-J_0(a)} \Lambda^{-1/2}(a)\, d\mathcal{H}^{n-k} a\ . \end{aligned}$$

which gives the desired convergence of measures $P_\lambda \Rightarrow P$.

Finally we consider the case $j > 1$ of the Main Theorem involving the extra disjoint regions $U_2, \cdots, U_j$ on each of which the Hessian $D^2 J_1$ has constant rank strictly larger than k_1, which is its rank on U_1. Now each set

$$M_i \ = \ \{x \in U_i \ : J_1(x) = 0\}$$

is, by Theorem 4.1(2), a compact $n - k_i$ dimensional submanifold. We will repeat most of the above arguments and again use the factor $\lambda^{k_1/2}$ to try to estimate

$$\lambda^{\frac{k_1}{2}} \int_{\mathbb{R}^n} \phi(x) e^{-J_0(x)-\lambda J_1(x)}\, dx\ .$$

as $\lambda \to \infty$. As before we may, by Lemma 5.4, restrict our integration to any fixed neighborhood U of M. We take $U = U_1 \cup \cdots \cup U_j$ where each U_i is, as before, a sufficiently small (depending on a given ε and test function ϕ) tubular neighborhood of M_i.

For the region U_1, we find, by estimating the upper and lower bounds just as before, that

$$\lim_{\lambda \to \infty} \lambda^{\frac{k_1}{2}} \int_{U_1} \phi(x) e^{-J_0(x) - \lambda J_1(x)}\, dx \;=\; 2^{\frac{k_1}{2}}\, k_1 \alpha_{k_1} \beta_{k_1} \int_{M_1} e^{-J_0(a)} \Lambda^{-1/2}(a)\, d\mathcal{H}^{n-k_1} a \tag{6.7}$$

where $\Lambda(a)$ is, as before, the product of the k_1 positive eigenvalues of $D^2 J_1(a)$ for $a \in M_1$.

However, for any region U_i with $i = 2, \cdots, j$, one finds that,with $a \in M_i$, in place of (6.4), one has the upper estimate

$$\begin{aligned}
&\lambda^{\frac{k_1}{2}} \int_{N_\delta(a)} e^{-\lambda J_1(G_a(y))}\, dy \\
&\leq\; \lambda^{\frac{k_1}{2}} \int_{N_\delta(a)} e^{-\frac{1}{2}\lambda \lambda_1(a)|y|^2)} (1 \;+\; 2C\lambda |y|^3)\, dy \\
&=\; \lambda^{\frac{k_1}{2}}\, k_i \alpha_{k_i} \int_0^\infty e^{-\frac{1}{2}\lambda\lambda_1(a) r^2} r^{k_i - 1}\, dr \;+\; 2C\lambda^{\frac{k_1+2}{2}}\, k_i \alpha_{k_i} \int_0^\infty e^{-\frac{1}{2}\lambda\lambda_1(a) r^2} r^{3+k_i-1}\, dr \\
&\longrightarrow\; 0 \;+\; 0 \quad \text{as } \lambda \to \infty\,,
\end{aligned} \tag{6.8}$$

by (5.1), because $k_i > k_1$ and $3 + k_i > k_1 + 2$. It follows that, for any bounded continuous ϕ on $\mathbb{R}^n$,

$$\lim_{\lambda \to \infty} \lambda^{\frac{k_1}{2}} \int_{U_i} \phi(x) e^{-J_0(x) - \lambda J_1(x)}\, dx \;=\; 0 \tag{6.9}$$

for $i = 2, \cdots, j$. With

$$Y_\lambda \;=\; \left(2^{\frac{k_1}{2}}\, k_1 \alpha_{k_1} \beta_{k_1}\right)^{-1} Z \lambda^{\frac{k_1}{2}}\,,$$

we conclude from (5.1), (6.7), and (6.9) as before that, as $\lambda \to \infty$, $Y_\lambda / Z_\lambda \to 1$ and

$$Z_\lambda \int_{\mathbb{R}^n} \phi(x) e^{-J_0(x) - \lambda J_1(x)}\, dx \;\to\; Z \int_{M_1} \phi(a) e^{-J_0(a)} \Lambda^{-1/2}(a)\, d\mathcal{H}^{n-k_1} a\,,$$

which completes the proof.

7. Multiple Limits

Multiple constraints generated by functions $J_1, J_2, \cdots, J_j$ lead to consideration of Gibb's measures obtained from multiple limits $\lim_{\lambda_1 \to \infty} \lim_{\lambda_2 \to \infty} \cdots \lim_{\lambda_j \to \infty}$. With suitable nondegeneracy assumptions on the Hessians of the J_i, one expects to obtain consecutively suitably weighted Hausdorff measures on lower and lower dimensional submanifolds. To inductively follow this procedure, one first needs to verify that:

The Main Theorem remains true if $\mathbb{R}^n$ is replaced by an n dimensional Riemannian manifold N.

Proof. The proof follows exactly as in Section 6 until we need to apply Lemma 3.6 instead of Lemma 3.4 and replace the righthand side of (6) by

$$\lambda^{\frac{k}{2}} \int_M \left(\int_{V_\delta(a)} \phi(G_a(y)) \cdot e^{-J_0(G_a(y))} \cdot e^{-\lambda J_1(G_a(y))} \cdot (J_{n-k}\Pi_M)^{-1}(G_a(y)) \cdot J_k(G_a(y))\, dy \right) d\mathcal{H}^{n-k} a\,. \tag{7.1}$$

Note that, by our choice of U and (3.3), with $x = G_a(y) \in U$ (hence $a = \Pi_M(x)$), one also has

$$|y| \;\le\; \mathrm{dist}(x,a) \;<\; \varepsilon\,.$$

We estimate the first four factors of (7.1) as before in our estimate of (6). The new fifth factor is estimated using Lemma 3.3 to obtain

$$J_k(G_a(y)) \;\le\; (1 + C\varepsilon^\alpha)\,.$$

Combining these and changing C, we now have the upper bound

$$\begin{aligned} &\lambda^{\frac{k}{2}} \textstyle\int_U \phi(x) e^{-J_0(x) - \lambda J_1(x)}\, dx \\ &\le\; \lambda^{\frac{k}{2}} (1 + C\varepsilon^\alpha) \textstyle\int_M (\phi(a) + \varepsilon) e^{-J_0(a)} \int_{V_\delta(a)} e^{-\lambda J_1(G_a(y))}\, dy\, d\mathcal{H}^{n-k} a\,. \end{aligned}$$

which corresponds to (6.3). The remainder of the proof now follows precisely as before in Section 6. □

Using this, we may now establish one result concerning multiple limits.

Theorem 7.1. *Suppose that $p > 0$ and that, for $i = 0, \cdots, j$, J_i is a nonnegative $C^{3,\alpha}$ function on $\mathbb{R}^n$ satisfying $J_i(x) \ge |x|^p$ for $|x|$ sufficiently large. Suppose also that the set*

$$M \;=\; \{x \in \mathbb{R}^n : J_1(x) = J_2(x) = \cdots J_j(x) = 0\}$$

is nonempty and lies in a bounded open set U on which each Hessian $D^2 J_i$ is nonnegative definite with a constant rank k_i for $i = 1, \cdots, j$. We assume that, for $a \in M$, the total Hessian $D^2(J_1 + \cdots + J_j)(a)$ has rank $k \equiv k_1 + \cdots + k_j$. Hence M is a $C^{2,\alpha}$ smooth $n-k$ dimensional manifold. Finally we assume that the images of the individual Hessians $D^2 J_i(a)$ are orthogonal.

For $\lambda_1, \cdots, \lambda_j \ge 1$, let $P_{\lambda_1 \cdots \lambda_j}$ be the probability measure on $\mathbb{R}^n$, whose density is

$$Z_{\lambda_1 \cdots \lambda_j} e^{-J_0(x) - \lambda_1 J_1(x) - \cdots - J_j(x)}$$

with

$$Z^{-1}_{\lambda_1 \cdots \lambda_j} \;=\; \int_{\mathbb{R}^n} e^{-J_0(x) - \lambda_1 J_1(x) - \lambda_2 J_2(x) - \cdots - \lambda_j J_j(x)}\, dx.$$

Then

$$P_{\lambda_1 \cdots \lambda_j} \Rightarrow P \quad \textit{as} \quad \lambda_1, \cdots, \lambda_j \to \infty\,,\ \textit{regardless of the order},$$

where, for any Borel set $B \subset \mathbb{R}^n$,

$$P(B) \;=\; Z \int_{M \cap B} e^{-J_0(a)}\, \Lambda_1(a)^{-1/2} \cdots \Lambda_j(a)^{-1/2}\, d\mathcal{H}^{n-k}(a)$$

with

$$Z^{-1} \;=\; \int_M e^{-J_0}\, \Lambda_1^{-1/2} \cdots \Lambda_j^{-1/2}\, d\mathcal{H}^{n-k}\,,$$

where $\Lambda_i(a)$ is the product of the k_i positive eigenvalues of $D^2 J_i(a)$.

Proof. Following carefully the proof of the Main Theorem, we see that the only essential modification necessary for treating the multiple limits is the step involved with finding, for fixed $a \in M$, a rotation of coordinates, so that, on the fiber, $\Pi_W^{-1}\{a\}$, $J_1(x)$ now has the form

$$\mu_1 y_1^2 + \cdots + \mu_{k_1} y_{k_1}^2 + \cdots + \mu_{k_1+k_2} y_{k_1+k_2}^2 + \cdots + \mu_k y_k^2$$

where $y = (y_1, \cdots, y_n) = x - a$. The point here is that because the Hessians $D^2 J_1(a), \cdots, D^2 J_j(a)$ have images that are completely orthogonal, they may be simultaneously diagonalized. □

8. Application of Theorem 2.1 to microstructures of a s bistable martensitic alloy

The result proven in this communication pertains to computation of microstructure underlying effective properties of functional materials. For the sake of clarity and simplicity we restrict ourselves to 2×2 vectorial situation of a bistable martenstic alloy. Functional materials have multiple equilibrium configurations exhibiting microscale domain patterns. Such patterns can be modeled using gradients of weakly differentiable maps $u : \mathbb{R}^2 \to \mathbb{R}^2$ that, in some appropriate sense, satisfy the first order differential inclusions of the form

$$\begin{aligned} &\nabla u \in \bigcup_{i=1}^{2} SO(2)U_i, \qquad \text{a.e. in } \Omega \subset \mathbb{R}^2, \\ &\{U_i \,|\, i = 1, 2, \ldots, 2\} = \{RU_1R^T \,|\, R \in P\}, \\ &u(x) = g(x), \quad x \in \partial\Omega, \\ &\nabla g \in \text{Closure of the Interior of Rank-1 Convex Hull of} \bigcup_{i=1}^{2} SO(2)U_i. \end{aligned} \tag{8.1}$$

$\mathcal{P}$ is a point group of rotation matrices that maps a referential configuration back to itself. The group $SO(2)$ denotes matrix rotations in $\mathbb{R}^2$. Namely, if $Q \in SO(2)$ then

$$Q = \begin{pmatrix} \cos(\alpha) & \sin\alpha \\ -\sin\alpha & \cos(\alpha) \end{pmatrix}$$

for some angle $\alpha \in [0, 2\pi)$. The point group $\mathcal{P}$ is given by rotations that map any plane parallelopepid back to itself in the case $n = 2$ considered here. The variants $U_i \in M^{2\times 2}$ are assumed to be symmetric and positive definite Bain (transormation) matrices, describing the distortion of the atomic lattice.

In the framework of crystalline microstructures characterizing functional materials, minimizing stored energy oftentimes means to deal with the lack of lower semicontinuity. This means that the weak limit of a minimizing sequence does not minimize the associated stored energy. At the staring point let us assume that we are given a twice differentiable function $W : \mathbb{R}^{2\times 2} \mapsto \mathbb{R}^1$ such that, in view of (8.1),

$$\begin{aligned} &\{F \in \mathbb{R}^{2\times 2} \,|\, W(F) = 0\} = \bigcup_{i=1}^{2} SO(2)U_i, && && \text{positive otherwise}, \\ &W(QF) = W(F), && F \in \mathbb{R}^{2\times 2}, Q \in SO(2), && \text{frame indifference}, \\ &W(FR) = W(F), && F \in \mathbb{R}^{2\times 2}, R \in \mathcal{P}, && \text{material symmetry}. \end{aligned} \tag{8.2}$$

Then we assign to each matrix F, representing the deformation gradient, a number $I(u)$ – the strain energy, given by

$$I(u) \stackrel{\text{def}}{=} \int_{\Omega} W(\nabla u(x))\, dx. \tag{8.3}$$

The equilibria of I, i.e., the solutions of (8.1), are given by the variational problem

$$\inf \left\{ I(u) \,|\, u \in W^{1,p}(\Omega),\ u - g \in W_0^{1,p}(\Omega) \right\}. \tag{8.4}$$

Now, let u_j be a minimizing sequence of (8.4). Starting with a state with finite energy, such a sequence will be bounded. Then by Sobolev imbeddings, $2 < p < \infty$ or by Alaoglu(-Banach-Bourbaki) theorem, $p = \infty$, this sequence will posses either weak or weak-$*$ limit u. By the Fundamental Theorem of the $W^{1,p}$- gradient Young's measures, there exists a Radon probability measure, μ_x, describing the microstructure representing (possibly) non-attainable infima of I. Namely,

$$\lim_{j\to\infty} I(u_j) = \int_{\Omega} \int_{\mathcal{M}} W(A)\, d\mu_x(A)\, dx, \tag{8.5}$$

where $\mathcal{M} = M^{2\times 2}$. This implies

$$W^{effective}(x) = \int_{\mathcal{M}} W(A)\, d\mu_x(A).$$

The dependence of $W^{effective}$ on the spatial variable x is, in general, due to possible non-uniformity of the Young measure. Moreover, the effective stored energy, obtained by integration against the underlying Young's measure, coincides with the quasi-convexification of the crystalline density. The quasi-convexification is given by

$$I^{\#} \stackrel{\text{def}}{=} \inf \left\{ \frac{1}{\operatorname{meas}\Omega} \int\limits_{\Omega} W(Du)(x)\,dx \,|\, u \in W^{1,p}(\Omega),\, u(x) = g(x),\, x \in \partial\Omega \right\}$$
$$= \int\limits_{\Omega} W^{effective}(x)\; dx.$$

There does not exist a general method for computing $I^{\#}$ based on the above definition. Hence, the proposed sharpening of the initial Gibbs measure, by taking the limits $\lambda_i \to +\infty$ and $h \to 0_+$, yields approximation strategy, in the measure-theoretic framework, to identify $I^{\#}$.

The structure of the probability measure can be simple, if $\mu_x = \delta_{\nabla u(x)}$, or it can be profoundly complex. This depends on the boundary data g. The measure can be unique or it can suffer from a massive lack of non-uniqueness due to the complexity of the solutions to (8.1), [9]. Due to a possible lack of weak lower-semicontinuity, i.e., due to the possible lack of functional minimizers of (8.4), let us define

$$I(\mu) \stackrel{\text{def}}{=} \int\limits_{\Omega} \int_{\mathcal{M}} W(A)\,d\mu_x(A)\,dx, \tag{8.6}$$

and let

$$YM \stackrel{\text{def}}{=} \{\mu_x \,|\, \text{ probability measures generated by the gradients of bounded sequances in } \mathcal{A}\}$$

where $\mathcal{A} \stackrel{\text{def}}{=} \left\{u \in W^{1,p}(\Omega),\; u - g \in W_0^{1,p}(\Omega)\right\}$. The generalized variational principle reads

$$\inf \left\{ I(\mu) \,|\, \mu \in YM \right\}. \tag{8.7}$$

In the next sections, our goal is to show how the presented theory can be used to solve the generalized variational problem (8.7).

8.1. Reformulation of (8.4) using Lagrange multipliers within the presented theory

We reformulate the variational problem (8.4) using Lagrange multipliers to separate the various requirements. Namely, we have the following constraints

- C_1: $\nabla u \in \bigcup_{i=1}^{M} SO(2)U_i$, for a. a. $x \in \Omega$,
- C_2: continuity in view of the weak differentiability,
- C_3: attainment of the boundary condition.

Let us assume that each of the constraints admits a suitable "density", which we denote W_i, $i = 1, 2, 3$. Consequently, the global forms of the above constraints are given by $J_i(u) \stackrel{\text{def}}{=} \int_\Omega W_i(u(x))\,dx$. The densities are constructed in Section 8.2 and Section 8.3.

In order to re-phrase this problem in terms of our theory, we proceed as follows. Let $h > 0$ be a discretization parameter of any acceptable partitioning of Ω. We assume that the deformation u has a finite dimensional image given by a Finite Element approximation $\boldsymbol{u}_h$ that has coordinates $\{\{U_i^j\}_{i=1}^{N(h)}\}_{j=1}^2$ with respect to a suitable finite element basis $\{\varphi_i\}_{i=1}^{N(h)}$. Let us denote the space of such functions by V_h, i.e., $\boldsymbol{u}_h \in V_h$ has the representaion $\boldsymbol{u}_h^j(x) = \sum_{i=1}^{N(h)} U_i^j \varphi_i(x)$, $j = 1, 2$, and $\dim V_h = 2N(h)$. In what follows we will write $\boldsymbol{u}_h$, which is a function, instead of the vector $\{U_i\}_{i=1}^{2N(h)} = \left\{\{U_i^j\}_{i=1}^{N(h)}\right\}_{j=1}^2$. Let

$$M_h \stackrel{\text{def}}{=} \{\boldsymbol{u}_h \in V_h \,|\, J_1(\boldsymbol{u}_h(x)) = 0,\; J_2(\boldsymbol{u}_h(x)) = 0\},$$

be a subset of V_h of those discrete functions that satisfy the gradient and continuity constraints in Ω.

Remark 8.1. *We note that for M_h to be non-empty an adaptive partitioning of Ω is required in order to align the inter-element boundaries with the set of points at which the gradient of u_h suffers a discontinuity. The reason is that $u_h\big|_{\omega_h} \in C^\infty(\omega_h)$.* □

Then the finite dimensional version of the variational problem (8.4), may be written as a relaxed constrained optimization problem in the following form

$$\operatorname{Argmin}\{J_3(\boldsymbol{u}_h) \,|\, \boldsymbol{u}_h \in M_h\}.$$

We choose to replace the variational problem (8.4) in finite dimension with a relaxed (unconstrained) variational problem

$$\lim_{\lambda_2 \to +\infty} \lim_{\lambda_1 \to +\infty} \min_{\boldsymbol{u}_h \in V_h} I(\boldsymbol{u}_h; \lambda_1, \lambda_2).$$

where

$$I(\boldsymbol{u}_h; \lambda_1, \lambda_2) \stackrel{\text{def}}{=} \lambda_1 J_1(\boldsymbol{u}_h) + \lambda_2 J_2(\boldsymbol{u}_h) + J_3(\boldsymbol{u}_h). \tag{8.8}$$

We prove the following Lemma before we proceed to application Theorem 7.1.

Lemma 8.2. *The set M_h is a union of isolated orbits in $\mathbb{R}^{2N(h)}$.*

Proof. □

Let a Gibbs density f be given by

$$f\left(\boldsymbol{X}; \lambda_1, \lambda_2\right) \stackrel{\text{def}}{=} Z_{\lambda_1,\lambda_2} e^{-I\left(\boldsymbol{X}; \lambda_1, \lambda_2\right)},$$

$\boldsymbol{X} \in \mathbb{R}^{2N(h)}$, where Z_{λ_1,λ_2} is the normalizing constant, given by

$$Z^{-1}_{\lambda_1,\lambda_2} \stackrel{\text{def}}{=} \int_{\mathbb{R}^{2N(h)}} e^{-I\left(\boldsymbol{X}; \lambda_1, \lambda_2\right)} \, d\boldsymbol{X}.$$

We note that the Gibbs measure is, in general, an asymptotic solution of the Fokker-Planck equation. In our particular case, the standard variation σ^2 appearing in the Fokker-Planck equation, and, consequently, in the definition of the Gibbs measure, is set to be 2. Let us consider the Gibbs measure $P_{\lambda_1,\lambda_2,x,h}$ given by

$$P_{\lambda_1,\lambda_2,h}(B) = \int_B f\left(\boldsymbol{u}_h(x); \lambda_1, \lambda_2\right) \, dU_1, \ldots, dU_{2N(h)}, \tag{8.9}$$

where $B \subset \mathbb{R}^{2N(h)}$ is a Lebesgue measurable set in the phase space V_h containing the coordinates of $\boldsymbol{u}_h$. Consequently, the Gibbs measure is a measure on the discrete function space V_h. Assuming, for a moment, that the assumptions of Theorem 7.1 are satisfied, we obtain weak convergence of the family of Gibbs' probability measures $P_{\lambda_1,\lambda_2,h}$ to a Gibbs probability measure P_h, which is absolutely continuous with respect to a Hausdorff measure. Namely,

$$P_{\lambda_1,\lambda_2,h} \Rightarrow P_h, \qquad \text{as } \lambda_1 \to +\infty, \lambda_2 \to +\infty. \tag{8.10}$$

In addition, the order in which the limits are taken can be arbitrary for we assume that the images of the individual Hessians of J_1 and J_2 are orthogonal when restricted to the relaxing orbits $SO(2)U_1 \cup SO(2)U_2$. Next, we need to identify the dimensionality of the Hausdorff measure. In view of (8.9), we consider the map

$$\left\{\{U_j\}_{j=1}^{2N(h)}\right\} \mapsto J_1(\nabla u_h) + J_2(u_h).$$

The Hessian corresponding to this maping is $4N(h) \times 4N(h)$ symmetric matrix. Let us assume that the rank of the Hessian of $J_1 + J_2$ restricted to the relaxing orbits is equal to $2N(h)$, i.e., the Hessian has a zero eigenvalue of multiplicity $2N(h)$ when evaluated at $SO(2)U_1 \cup SO(2)U_2$. Theorem 7.1 identifies the dominating measure to be $(n-k)$-dimensional Hausdorff measure. It follows from Lemma 8.2 that the dominating measure is supported on the union of orbits in $\mathbb{R}^{2N(h)}$. Thus it has to be a one-dimensional Hausdorff measure. In our case, $n = 2N(h)$, hence $k = 2N(h) - 1$. For any Borel set $B \subset \mathbb{R}^{2N(h)}$, the resulting Gibbs probability measure is thus given by

$$P_h(B) = Z \int_{M_h \cap B} e^{-J_3(v_h)} \Lambda_1(v_h)^{-1/2} \Lambda_2(v_h)^{-1/2}\, d\mathcal{H}^1(v_h),$$
$$Z^{-1} = \int_{M_h} e^{-J_3(v_h)} \Lambda_1(v_h)^{-1/2} \Lambda_2(v_h)^{-1/2}\, d\mathcal{H}^1(v_h),$$

where, we recall, Λ_i are products, one for each relaxing orbit, of the $2N(h)$ positive eigenvalues of the second tensor derivative of $J_1 + J_2|_{v_h \in M_h}$, taken with respect to the coordinates $\{U_j\}_{j=1}^{2N(h)}$.

Let

$$\mathcal{L}_h \stackrel{\text{def}}{=} \lim_{\lambda_2 \to +\infty} \lim_{\lambda_1 \to +\infty} \operatorname*{Argmin}_{\boldsymbol{u}_h \in V_h} I(\boldsymbol{u}_h; \lambda_1, \lambda_2).$$

In words, we are collecting in $\mathcal{L}_h$ all continuous, discrete functions (more precisely their coordinates in V_h) with gradients in $SO(2)U_1 \cup SO(2)U_2$, upon suitable repartitioning of Ω, that minimize the $L^2(\partial\Omega)-$distance to a given boundary data (constraint) g. Now, suppose that we construct the Markov Chains of coordinates minimizing J_3 as $\lambda_i \to +\infty$, $i = 1, 2$. Then we have

Lemma 8.3. *Let*

$$M_h^{opt} \stackrel{def}{=} \operatorname{Argmin}\{J_3(\boldsymbol{u}_h) \,|\, \boldsymbol{u}_h \in V_h, J_1(\boldsymbol{u}_h(x)) = 0,\ J_2(\boldsymbol{u}_h(x)) = 0\}.$$

Then the set M_h^{opt} is a union of isolated points, and

$$\operatorname{card} M_h^{opt} \leq .$$

Thus, in particular, the dominating measure for the limiting Gibbs measure P_h is the zero-dimensional Hausdorff measure.

Proof. □

Applying Theorem 7.1 to Markov Chains of coordinates, which are optimal with respect to J_3, we obtain a limiting Gibbs measure, that, in view of Lemma 8.3, has now the representation

$$P_h(B) = Z^{-1} \sum_{v_h^* \in \mathcal{L}_h} q(v_h^*) \delta_{v_h^*}(B), \qquad \text{where} \qquad q(v_h^*) = e^{-J_3(v_h^*)} \Lambda_1(v_h^*)^{-1/2} \Lambda_2(v_h^*)^{-1/2},$$
$$Z^{-1} \sum_{v_h^* \in \mathcal{L}_h} q(v_h^*) = 1.$$

Remark 8.4. *We recall, that the convergence* (8.10) *takes place along the Markov Chains $\{U_{i;\lambda_1,\lambda_2}\}_{i=1}^{2N(h)}$ in V_h. The Gibbs measure on the discrete function space V_h induces a measure on the physical domain Ω. Using these coordinates, we construct maps $x \mapsto \nabla \boldsymbol{u}_{\lambda_1,\lambda_2,h}(x)$. Consequently, we construct a family of Radon measures, $\mu_{x,h}$, parameterized by $x \in \Omega$, characterizing volume fractions*

$$\lambda_{i,h}(x_0) \stackrel{def}{=} \lim_{R \to 0_+} \lim_{r \to 0_+} \{x \in B_R(x_0) \,|\, \operatorname{dist}\{\nabla \boldsymbol{u}_h(x), SO(2)U_i\} < r\} \,/\, \operatorname{meas}(B_R(x_0)), \qquad i = 1, 2, \tag{8.11}$$

for any $x_0 \in \Omega$. Hence, there exists a family of linear bounded operators $T_x : P_{\lambda_1,\lambda_2,h} \mapsto \mu_{\lambda_1,\lambda_2,x,R,r,h}$. Then, in view of Theorem 7.1,

$$\mu_{\lambda_1,\lambda_2,x,R,r,h} \Rightarrow \mu_{x,R,h}, \qquad \text{as } \lambda_1 \to +\infty, \lambda_2 \to +\infty.$$

We note that the "volume averaging" (8.11) *of the gradients of the induced sequences is not the only way to compute the volume fractions. In terms of the Gibbs measure, it is more appropriate to compute the volume fractions as follows.*

In this paper, we do not investigate the limit $h \to 0_+$. Nonetheless, we expect

$$\mu_{x,R,h} \Rightarrow \mu_x \qquad \text{as } R, h \to 0_+.$$

In connection with the generalized variational principle (8.7)*, we have in the sense of the weak convergence of measures, regardless of the order in which the limits are taken,*

$$\lim_{h \to 0_+} \lim_{\lambda_2 \to +\infty} \lim_{\lambda_1 \to +\infty} \mu_{\lambda_1,\lambda_2,x,h} = \operatorname{Arginf}\{E(\mu) \,|\, \mu \in YM\}.$$

We will provide a communication of our attempts to prove these results in a separate exposition. □

The next two sections contain a particular construction of the densities for J_1 and J_2, which satisfy the assumptions of Theorem 7.1 with $k_i = 2$ for $i = 1, 2$.

8.2. A VECTORIAL GRADIENT CONSTRAINT DENSITY FOR A BINARY ALLOY

A proper gradient constraint density has to satisfy the constrains (A5) and (A6), in addition to the requirements (8.2), appearing as the assumptions in Lemma 2.1, in order to avoid spurious states that could pollute the resulting measure as the Lagrange multipliers approach infinity. We restrict ourselves to a two dimensional vectorial case for simplicity.

Let us consider the following two symmetric, positive definite, Bain (transformation) matrices

$$U_1 \stackrel{\text{def}}{=} \begin{pmatrix} \alpha & \gamma \\ \gamma & 1 \end{pmatrix} \qquad \text{and} \qquad U_2 \stackrel{\text{def}}{=} \begin{pmatrix} \beta & \gamma \\ \gamma & 1 \end{pmatrix}, \tag{8.12}$$

where $\alpha, \beta, \gamma \in \mathbb{R}$, $\alpha \neq \beta$, $\alpha \neq \gamma^2$, $\beta \neq \gamma^2$, are assumed to be given. We note that $\operatorname{rank}(U_1 - U_2) = 1$. We have the following

Lemma 8.5. *Let $a_i \in R^+$, $i = 1, \ldots, 5$, let e_i, $i = 1, 2$ be canonical basis vectors in $\mathbb{R}^2$, $e \stackrel{def}{=} (1,1)^T$, and let us define a gradient constraint density by*

$$\begin{aligned} W_1(F) \stackrel{def}{=} \;& a_1 \left(\det F - \left(\alpha - \gamma^2\right)\right)^2 \left(\det F - \left(\beta - \gamma^2\right)\right)^2 \\ &+ a_2 \left(|\mathrm{Cof}\, F e_2|^2 - \left(\alpha^2 + \gamma^2\right)\right)^2 \left(|\mathrm{Cof}\, F e_2|^2 - \left(\beta^2 + \gamma^2\right)\right)^2 \\ &+ a_3 \left(|F e_1|^2 - (\alpha^2 + \gamma^2)\right)^2 \left(|F e_1|^2 - (\beta + \gamma^2)\right)^2 \\ &+ a_4 \left(|F e_2|^2 - (\gamma^2 + 1)\right)^2 \\ &+ a_5 \left(|F e|^2 - (\alpha + \gamma)^2 - (\gamma + 1)^2)\right)^2 \left(|F e|^2 - (\beta + \gamma)^2 - (\gamma + 1)^2\right)^2 . \end{aligned} \tag{8.13}$$

Then for any $a_i \in R^+$, $i = 1, \ldots, 5$, the strain density constraint W_1 has the properties listed below.

(1) *It satisfies the conditions* (8.2).
(2) *It satisfies the conditions (A5) of Theorem 2.1, with*

$$\operatorname{rank} \int_\Omega D^2 W_1(\nabla \boldsymbol{u}_h(x))|_{\nabla \boldsymbol{u}_h(x) \in SO(2)U_1 \cup SO(2)U_2} \, dx = 2N_h \times 2N_h - b_h,$$

where b_h is the number of the "boundary degrees of freedom". This number is defined below.
(3) *It satisfies the condition (A6) of Theorem 2.1.*

Proof. Verification of (8.2)*b.* This property follows immediately from the fact that $\det(QF) = \det F$ and, using $\mathrm{Cof}\, F = \det F F^{-T}$, [1] we obtain for any $v \in \mathbb{R}^2$

$$|\mathrm{Cof}(QF)v|^2 = (\det F)^2 \left(QF^{-T} v\right)^T QF^{-T} v = (\det F)^2\, v^T F^{-1} F^{-T} v = |\mathrm{Cof}(F)v|^2 .$$

The last three contributions in the definition (8.13) contain only the terms $|Fe_i|^2$, $i = 1, 2, 3$, $e_3 = (1,1)^T$, which are unchanged by any unitary rotation.

Verification of (8.2)*c.* This can be proven identically to the previous step.

[1] Notice that $\det F = \frac{1}{n} F : (Cof F)^T$ for $F \in M^{n \times n}$, where ":" denotes Frobenius matrix multiplication given for two compatible matrices A abn B by $A : B = \mathrm{Tr}(AB)$.

Verification of (8.2)*a.* Since $W_1(U_1) = W_1(U_2) = 0$, the above proof of the frame indifference, property (8.2b), shows that any matrix $F \in SO(2)U_1 \bigcup SO(2)U_2$ belongs to the null set of W_1. It remains to verify the opposite inclusion. Hence, let $W_1(F_0) = 0$, where

$$F_0 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Then we get the following set of equations

$$\begin{aligned} ad - b^2 &= \alpha - \gamma^2 \quad \text{or} \quad = \beta - \gamma^2 \\ a^2 + b^2 &= \alpha^2 + \gamma^2 \quad \text{or} \quad = \beta^2 + \gamma^2 \\ a^2 + c^2 &= \alpha^2 + \gamma^2 \quad \text{or} \quad = \beta^2 + \gamma^2 \\ b^2 + d^2 &= \gamma^2 + 1 \\ (a-b)^2 + (c-d)^2 &= (\alpha - \gamma)^2 + (\gamma - 1)^2 \quad \text{or} \quad = (\beta - \gamma)^2 + (\gamma - 1)^2. \end{aligned} \tag{8.14}$$

Solving with the right-hand sides given by $\alpha - \gamma^2$, $\alpha^2 + \gamma^2$ and $(\alpha - \gamma)^2 + (\gamma - 1)^2$, we obtain two solutions to (8.14). Namely,

$$(a, b, c, d) \in \{(-\alpha, -\gamma, -\gamma, -1), (\alpha, \gamma, \gamma, 1)\}.$$

These solutions are related by $SO(2)$ rotations. Namely, by the Identity matrix, and by the negative Identity matrix, which corresponds to the rotation matrix Q with $\alpha = \pi$. Hence, we conclude that these two solutions are in $SO(2)U_1$. Similarly, solving (8.14) with the right-hand sides containing β, we obtain two solutions in $SO(2)U_2$. If we intermix the right-hand sides there is no solution to (8.14).

Verification of the condition (A5) of Theorem 2.1. Let us denote by d the tensor derivative with respect to the coordinates $\{U_i\}_{i=1}^{2N(h)}$, and let us denote by D the tensor derivative with respect to the components of the deformation gradient. First we observe that

$$\nabla u_h(x) \in SO(2)U_1 \cup SO(2)U_2 \implies -\operatorname{div} DW_1(\nabla u_h(x)) = 0 \qquad \text{for a.a. } x \in \Omega, \tag{8.15}$$

which is the basic equation of elasticity since $DW(\cdot)$ is the stress tensor represented by a 2×2 matrix. We show below that, $d_{kl} \stackrel{\text{def}}{=} \partial / \partial U_k^l$,

$$\left(\int_\Omega d^2 W_1(\nabla u_h(x))\, dx \right)_{kl} = d_{kl} \int_\Omega DW_1(\nabla u_h(x)) \nabla \varphi(x)\, dx. \tag{8.16}$$

Hence, using (8.15) and integration by parts, we obtain

$$\left(\int_\Omega d^2 W_1(\nabla u_h(x))|_{SO(2)U_1 \cup SO(2)}\, dx \right)_{kl} = \int_{\partial\Omega} d_{kl} DW_1(\nabla u_h(x))|_{SO(2)U_1 \cup SO(2)} \cdot n \varphi(x)\, dS. \tag{8.17}$$

Now, we perform the d – tensor derivative of this equation. We obtain

$$-d_{kl} \operatorname{div} DW_1(\nabla u_h(x)) = -\operatorname{div} d_{kl} DW_1(\nabla u_h(x)) = -\operatorname{div} \sum_{p=1}^{2} \sum_{q=1}^{2} \partial_{pl} \partial_{qk} W_1(\nabla u_h) \frac{\partial \varphi_l^p}{\partial x_p} \frac{\partial \varphi_l^k}{\partial x_q}. \tag{8.18}$$

□

8.2.1. A continuity constraint density

Suppose that Ω is partitioned into $\omega_1 \cup \omega_2 \cup \ldots \cup \omega_k$. Let $\boldsymbol{u}_h \in V_h$. In order to impose the continuity across the interelement boundaries, we define

$$W_2(\boldsymbol{u}_h(x_0)) \stackrel{\text{def}}{=} \frac{1}{2} [\![\boldsymbol{u}_h(x_0)]\!]^2 \stackrel{\text{def}}{=} \frac{1}{2} \left\| \lim_{\substack{x \to x_0 \\ x \in \omega_i}} \boldsymbol{u}_h(x) - \lim_{\substack{x \to x_0 \\ x \in \omega_j}} \boldsymbol{u}_h(x) \right\|^2, \quad x_0 \in \Omega.$$

We note that $W_2(\boldsymbol{u}_h(x)) = 0$ for any x in the interior of any ω_i. Then

$$J_2(\boldsymbol{u}_h) = \sum_{i,j} \int_{\partial\omega_i \cap \partial\omega_j} W_2(\boldsymbol{u}_h(s))\, dS$$

vanishes if and only if the continuity constraint is satisfied, i.e., when the condition C_2 holds. We note that the Hessian of W_2 is the zero matrix, i.e., all entries in D^2W are zero, when restricted to M_h, since $u_h \in M_h$ implies $u_h \in C^0(\Omega)$. Consequently,

(1) $\operatorname{rank}\left(D^2W_1 + D^2W_2\right)|_{M_h} = \operatorname{rank}\left(\boldsymbol{D}^2W_1\right)|_{M_h} = 2N(h)$,
(2) $\operatorname{Tr}\left(\left(D^2W_1\right)^T D^2W_2\right)|_{M_h} = 0$,
(3) Consequently, it follows that M_h is a zero-dimensional submanifold of $\mathbb{R}^{4N(h)\times 4N(h)}$. In other words, M_h contains only isolated points, c. f. Lemma 8.2.

8.3. A BOUNDARY CONSTRAINT DENSITY

Taking $W_3(\boldsymbol{u}(s)) = \frac{1}{2}\left\|\boldsymbol{u}(s) - \boldsymbol{g}(s)\right\|^2$, $s \in \partial\Omega$, we have that

$$J_3(\boldsymbol{u}) = \int_{\partial\Omega} W_3(\boldsymbol{u}(s))\, dS$$

is minimized precisely when the boundary constraint holds, i.e., when the constraint C_3 is satisfied.

9. CONCLUSIONS

We here propose some alternatives to the Markov Chain Monte Carlo approach of the previous section. One straightforward alternative is to perform the MCMC on a set of approximants which already satisfy one of the constraints. The most obvious version of this approach is to only consider u which already satisfy the gradient constraint C_3. It is easy to generate random variates from probability distributions on $SO(N)$ and to use the product of Haar measures as the dominating measure for the proposal density $r(v, w)$. It is also necessary to be able to jump between "wells", meaning the cossets $SO(N)U_i$, $1 \le i \le M$. This seems straightforward, but it will probably be necessary to introduce further parameters, e.g., stay in the current well with probability q, and otherwise jump to random well with probability $(1-q)/(M-1)$, selecting from $SO(n)$ using the Haar probability measure. For movements within a well, we would naturally use proposal distributions absolutely continuous with respect to Haar measure. We will look for distributions which are easy to use and give good convergence properties.

It seems possible that we could enforce the continuity constraint, but this looks very difficult.

Another strategy which has some promise is the so-called Gibbs sampler version of MCMC [12]. The idea here is to update each component of v one at a time in some sweep through the components. In order to keep the Gibbs distribution as the stationary distribution of the Markov Chain, it is necessary to use the conditional distribution of the component (say) v_i given all the other components $(v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k)$. This is determined from the (unnormalized) density for Gibbs distribution

$$\exp\left(-I(v,\lambda)/(2\sigma^2)\right)$$

as a function of v_i. Very efficient algorithms have been developed for sampling from general densities given in exponential form (see [11]). There are numerous variants of this approach – e.g. instead of systematically sweeping through the components v_i, one can randomly sample them as well.

10. ACKNOWLEDGEMENTS

Dennis Cox was supported in part by grant NSF DMS-0204723.
Robert Hardt was supported in part by grant NSF DMS-0306294.

Petr Klouček was supported in part by the grants NSF ACI-0325081, the grant NSF CCR-0306503, and by the European Commission via MEXC-CT-2005-023843.

REFERENCES

1. P. Billingsley, *Probability and measure*, 1 ed., John Wiley and Sons, 1995.
2. I. Bohachevsky, M. Johnson, and M. Stein, *Generalized simulated annealing for function optimization*, Technometrics **28** (1986), 209–217.
3. A. Corana, M. Marchesi, C. Martini, , and S. Ridella, *Continuous variables with the simulated annealing algorithm*, ACM Transactions on Mathematical Software **13** (1987), 262–280.
4. J. Dieudonné, *Foundations of modern analysis*, 1 ed., Academic Press, New York, 1960.
5. H. Federer, *Curvatures measures*, Trans. AMS **93** (1959), 418–491.
6. ______, *Geometric measure theory*, Springer-Verlag, New York, 1969.
7. M. Locatelli, *Simulated annealing algorithms for continuous global optimization: Convergence conditions*, Journal of Optimization Theory and Applications **104** (2000), 121–133.
8. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, and Teller E, *Equation of state calculation by fast computing machines*, Journal of Chemical Physics **21** (1953), 1087–1091.
9. P. Pedregal, *Variational methods in nonlinear elasticity*, SIAM, Philadelphia, 2000.
10. Kirkpatrick S, Gelatt CD, and Vecchi MP, *Optimization by simulated annealing*, Science **220** (1983), 671–680.
11. Gilks WR and Wild P, *Adaptive rejection sampling for gibbs sampling*, Applied Statistics **41** (1992), 337–348.
12. Gilks WR, Richardson S, and Spiegelhalter DJ, *Markov chain monte carlo in practice*, Chapman-Hall/CRC, 1995.