JOURNAL OF APPROXIMATION THEORY 1, 444–451 (1968)

# Convergence of Metric Nuclei

JAY E. ISRAEL

*Department of Mathematics, University of Western Australia, Nedlands, W.A., Australia*

## 1. INTRODUCTION

A subset of a metric space is *totally bounded* if, for every $\epsilon > 0$, it can be covered by a finite number of sets, each with diameter $\leqslant 2\epsilon$. If $A$ is totally bounded, $N_\epsilon(A)$ is the smallest number of sets in such a covering. Using $N_\epsilon(A)$ as a coarse indicator of the size of a set, G. Strang introduced the notion that a set can have a nucleus, a pre-eminent element at which the mass of the set is concentrated (see [*1*]). Using the terminology of [*2*], we say that $x$ is the *nucleus* of $A$, if, for every neighbourhood, $V$, of $x$,

$$N_\epsilon(A \cap V) \sim N_\epsilon(A) \qquad (\epsilon \to 0).$$

In heuristic terms, $x$ is the nucleus of $A$, if every neighbourhood of $x$ in $A$ is asymptotically as large as all of $A$. Fix $L > 0$ and $T > 0$. Let $0 = t_0 < t_1 < \ldots < t_{m-1} < t_m = T$ and let $c_1, c_2, \ldots, c_m$ be fixed constants with $c_m \neq 0$. Using the uniform metric ($d(F,G) = \sup_{0 \leqslant t \leqslant T} |F(t) - G(t)|$), define $\Lambda_L$ to be the set of real-valued functions, $F$, on $[0,T]$, such that $F(0) = 0$ and, for all $x,y \in [0,T]$, $|F(x) - F(y)| \leqslant L|x-y|$ (i.e. $F$ satisfies a Lipschitz condition with coefficient $L$). If

$$A = \left\{ F \in \Lambda_L \colon \sum_{i=1}^{m} c_i F(t_i) = 1 \right\}$$

is nonvoid (in Section 3 we shall derive a sufficient condition for $A$ to be nonvoid), then the following is known to hold (see [*2*]):

1.1. THEOREM. *A has a nucleus the broken line joining the points* $(0,\theta_0), (t_1,\theta_1), \ldots, (t_m,\theta_m)$, *where* $\theta_0 = 0$ *and* $\theta_1, \ldots, \theta_m$ *are the unique values which render the quantity*

$$B(\theta_1, \ldots, \theta_m) = \sum_{i=1}^{m} \{(L(t_i - t_{i-1}) + \theta_i - \theta_{i-1}) \ln (L(t_i - t_{i-1}) + \theta_i - \theta_{i-1})$$
$$+ (L(t_i - t_{i-1}) - \theta_i + \theta_{i-1}) \ln (L(t_i - t_{i-1}) - \theta_i + \theta_{i-1})\}$$

*a minimum, subject to* $\sum_{i=1}^{m} c_i \theta_i = 1$. (*We denote by* ln *the natural logarithm.*)

In this way, we obtain a nucleus by imposing the sum constraint $\sum_{i=1}^{m} c_i F(t_i) = 1$. It will be shown in Section 3 that this nucleus maximizes a certain functional over $A$. The question arises: what is the behaviour of the sequence of nuclei arising from a sequence of such sum constraints when the sums are Riemann sums for an integral $\int_0^T c(t) F(t)\, dt$? It will be shown in Section 4 that the sequence of nuclei converges to a function, $\tilde{F}$, which satisfies the integral constraint $\int_0^T c(t) \tilde{F}(t)\, dt = 1$. Furthermore, $\tilde{F}$ satisfies the maximization condition to be proven for the sum case. The question of whether or not $\tilde{F}$ is the nucleus of the set of functions in $\Lambda_L$ satisfying the integral constraint is still open.

## 2. Variational Results

This section contains definitions and technical results to be used in later convergence proofs. Here, we work on an arbitrary real interval, $[a, b]$.

2.1. Proposition. *If $p$ is an integrable function on $[a, b]$, then*

$$\lim_{\lambda \to \infty} \int_a^b p(t) \tanh(\lambda p(t))\, dt = \int_a^b |p(t)|\, dt.$$

*Proof.* Fix $\epsilon > 0$ and let $I_1 = \{t \in [a, b] : |p(t)| < \epsilon M/3\}$, where $M = (b - a)^{-1}$. By integrability, there is a $K$ such that, if $I_2 = \{t : |p(t)| > K\}$, then $\int_{I_2} |p(t)|\, dt < \epsilon/3$. For $\lambda > 0$, let us look at

$$\begin{aligned} I(\lambda) &= \left| \int_a^b p(t) \tanh(\lambda p(t))\, dt - \int_a^b |p(t)|\, dt \right| \\ &= \int_a^b |p(t)| \{1 - \tanh(\lambda |p(t)|)\}\, dt \\ &\leqslant \int_{I_1} + \int_{I_2} + \int_{I_3}, \qquad \text{where } I_3 = [a, b] - I_2 - I_1 \\ &\leqslant (b - a)\, \epsilon M/3 + \epsilon/3 + \int_{I_3} \\ &\leqslant 2\epsilon/3 + K \int_{I_3} \{1 - \tanh(\lambda \epsilon M/3)\}\, dt. \end{aligned}$$

Now let

$$\lambda_0 = 3(M\epsilon)^{-1} \operatorname{arctanh} \max\left\{0, \left(1 - \frac{M\epsilon}{3K}\right)\right\}.$$

Then if $\lambda > \lambda_0$,

$$\begin{aligned} I(\lambda) &\leqslant 2\epsilon/3 + K \int_{I_3} \left\{1 - \left(1 - \frac{M\epsilon}{3K}\right)\right\} dt \\ &\leqslant 2\epsilon/3 + K \frac{M\epsilon}{3K}(b - a) = \epsilon. \end{aligned}$$

Q.E.D.

Denote by $P_L$ the set of all Riemann integrable functions, $p$, on $[a,b]$, for which

$$\int_a^b |p(t)|\,dt > 1/L. \qquad 2.2$$

2.3. Corollary. *Let $p \in P_L$. Then there is a $\lambda > 0$ such that $\int_a^b p(t)\tilde{f}(t)\,dt = 1$, where $\tilde{f}(t) = L\tanh(\lambda p(t))$.*

*Proof.* $r(\lambda) = L\int_a^b p(t)\tanh(\lambda p(t))\,dt$ is a continuous function of $\lambda$ with $r(0) = 0$ and (by 2.1 and 2.2) $\lim_{\lambda\to\infty} r(\lambda) > 1$. The required result then follows from the intermediate value theorem.

Q.E.D.

If $p \in P_L$, define $D_L{}^p$ to be the set of Riemann integrable functions, $f$, satisfying $|f| \leqslant L$ a.e. and $\int_a^b p(t)f(t)\,dt = 1$. Then 2.3 shows that $\tilde{f} \in D_L{}^p$, so that $D_L{}^p$ is nonvoid. For $f \in D_L{}^p$, define

$$H(f) = -\int_a^b \left\{\frac{L+f(t)}{2L}\ln\frac{L+f(t)}{2L} + \frac{L-f(t)}{2L}\ln\frac{L-f(t)}{2L}\right\}dt.$$

2.4. Theorem. *Let $\tilde{f}$ be as in 2.3. If $f \in D_L{}^p$ and $f \neq \tilde{f}$ on a set of positive measure, then $H(f) < H(\tilde{f})$.*

*Proof.* Let $I(w) = H(\tilde{f} + w(f-\tilde{f}))$ for $0 \leqslant w \leqslant 1$. Then

$$I'(w) = -\frac{1}{2L}\int_a^b (f-\tilde{f})\ln\frac{L+\tilde{f}+w(f-\tilde{f})}{L-\tilde{f}-w(f-\tilde{f})}\,dt,$$

and

$$\begin{aligned} I'(0) &= -\frac{1}{2L}\int_a^b (f-\tilde{f})\ln\frac{1+\tanh(\lambda p)}{1-\tanh(\lambda p)}\,dt \\ &= -\frac{\lambda}{L}\int_a^b \{f(t)-\tilde{f}(t)\}p(t)\,dt \\ &= 0, \quad \text{since } f \text{ and } \tilde{f} \text{ are in } D_L{}^p. \end{aligned}$$

Furthermore,

$$I''(w) = -\int_a^b \frac{(f-\tilde{f})^2}{L^2 - \{\tilde{f}+w(f-\tilde{f})\}^2}\,dt < 0 \qquad (0 \leqslant w < 1)$$

since $|\tilde{f} + w(f-\tilde{f})| \leqslant (1-w)|\tilde{f}| + w|f| < L$ a.e. By the mean value theorem, then, $I(1) < I(0)$.

Q.E.D.

2.5. COROLLARY. *Let $\alpha$ and $\beta$ be fixed with $|\alpha - \beta| \leqslant L(b-a)$, and let $\tilde{F}(t)$ be the straight line joining the points $(a,\alpha)$ and $(b,\beta)$. If $F \neq \tilde{F}$ is any other function joining these two points and satisfying a Lipschitz condition with coefficient $L$, then $H(F') < H(\tilde{F}')$.*

*Proof.* $F'$ and $\tilde{F}'$ are in $D_L{}^p$ when $p(t) \equiv (\beta - \alpha)^{-1}$. By 2.4, $H$ is maximized in $D_L{}^p$ for $\tilde{f}(t) = L\tanh(\lambda p(t)) \equiv$ constant. Therefore, $H(F')$ is maximized when and only when $F$ is linear, i.e., $F = \tilde{F}$.

Q.E.D.

Let us now work on the interval $[0,T]$, where, as before, $T > 0$ is fixed. Denote by $C_L$ the set of Riemann integrable functions, $c$, on $[0,T]$, satisfying

$$c(T) \neq 0, \tag{2.6}$$

and

$$\int_0^T \left| \int_t^T c(s)\,ds \right| dt > 1/L. \tag{2.7}$$

Let $A = \{F \in \Lambda_L : \int_0^T c(t)\,F(t)\,dt = 1\}$. If $F \in A$,

$$\int_0^T c(t)\,F(t)\,dt = \int_0^T p(t)\,F'(t)\,dt, \tag{2.8}$$

where $p(t) = \int_t^T c(s)\,ds \in P_L$. Thus we have:

2.9. PROPOSITION. *If $F \in A$ then $F' \in D_L{}^p$. Conversely, if $f \in D_L{}^p$ then $F(t) = \int_0^t f(s)\,ds \in A$. Furthermore, $A$ is nonvoid, since $\tilde{f}$ (defined as in 2.3) is in $D_L{}^p$. Define $\tilde{F}(t) = \int_0^t \tilde{f}(s)\,ds$.*

2.10. PROPOSITION. *If $F \in A$ and $F \neq \tilde{F}$, then $H(F') < H(\tilde{F}')$.*

*Proof.* Immediate from 2.4.

## 3. CONNECTION WITH NUCLEI

Let us now relate maximization of $H$, as discussed in the previous section, to the notion of a nucleus.

3.1. THEOREM. *Let $A$ be as in* 1.1, *and let $\tilde{F}$ be its nucleus. Then $\tilde{F}$ is the unique element which maximizes $H(\tilde{F}')$ over $A$.*

*Proof.* As a consequence of 2.5, it is enough to consider functions $y_{x_1, x_2, \ldots, x_m}(t)$ which are broken line functions connecting the points $(0,0)$, $(t_1, x_1), \ldots, (t_m, x_m)$, where $|x_i - x_{i-1}| \leqslant L(t_i - t_{i-1})$ $(i = 1, 2, \ldots, m)$ and

$\sum_{i=1}^{m} c_i x_i = 1$. We must maximize $H(y'_{x_1, x_2, \ldots, x_m})$ over all admissible values of $x_1, \ldots, x_m$. But

$$H(y'_{x_1,\ldots,x_m}) = -\sum_{i=1}^{m} \int_{t_{i-1}}^{t_i} \{ \qquad \} dt$$

$$= -\frac{1}{2L} \sum_{i=1}^{m} \{(L(t_i - t_{i-1}) + x_i - x_{i-1}) \ln (L(t_i - t_{i-1}) + x_i - x_{i-1})$$

$$+ (L(t_i - t_{i-1}) - x_i + x_{i-1}) \ln (L(t_i - t_{i-1}) - x_i + x_{i-1})$$

$$- 2(t_i - t_{i-1}) \ln (2L(t_i - t_{i-1}))\}$$

$$= -\frac{1}{2L} B(x_1, \ldots, x_m) + \frac{1}{L} \sum_{i=1}^{m} (t_i - t_{i-1}) \ln (2L(t_i - t_{i-1})).$$

The last sum does not depend on the $x_i$. We conclude from 1.1, then, that $H(y'_{x_1,\ldots,x_m})$ is maximized when and only when $y_{x_1, x_2, \ldots, x_m} = \bar{F}$.

Q.E.D.

The quantity $H$ being maximized bears a striking resemblance to the "communication entropy" of information theory. What meaningful analogy (if any) can be drawn between the nucleus of a set and maximization of communication entropy is not clear, however.

In the previous section, we rewrote the integral $\int_0^T c(t) F(t) \, dt$ in the form $\int_0^T p(t) F'(t) \, dt$. Sum constraints can also be rewritten in this form, except that $p$ will not be continuous. The appropriate choice is

$$p(t) = \sum_{i=1}^{m} c_i u(t_i - t), \qquad \text{where } u(s) = \begin{cases} 1 & \text{when } s > 0 \\ 0 & \text{when } s \leqslant 0. \end{cases}$$

Indeed,

$$\int_0^T p(t) F'(t) \, dt = \sum_{i=1}^{m} \int_{t_{i-1}}^{t_i} \left\{ \sum_{j=i}^{m} c_j \right\} F'(t) \, dt$$

$$= \sum_{j=1}^{m} c_j \sum_{i=1}^{j} \int_{t_{i-1}}^{t_i} F'(t) \, dt$$

$$= \sum_{j=1}^{m} c_j F(t_j),$$

since $F(t_0) = 0$.

This allows us to conclude that $A$ is nonvoid if

$$\frac{1}{L} < \int_0^T |p(t)| \, dt = \sum_{j=1}^{m} \int_{t_{j-1}}^{t_j} \left| \sum_{i=j}^{m} c_i \right| dt.$$

that is, if

$$\sum_{j=1}^{m} \left| \sum_{i=j}^{m} c_i \right| (t_j - t_{j-1}) > \frac{1}{L}. \qquad 3.2$$

If 3.2 holds, we can write the nucleus in 1.1 explicitly as

$$\tilde{F}(t) = L \int_0^t \tanh\left\{\lambda \sum_{i=1}^{m} c_i u(t_i - s)\right\} ds$$

where $\lambda$ is chosen so that $\sum_{i=1}^{m} c_i \tilde{F}(t_i) = 1$ (possible by 2.3).

## 4. Convergence

Let $c \in C_L$ and let $p(t) = \int_t^T c(s)\,ds$. Let

$$A = \left\{F \in \Lambda_L : \int_0^T c(t) F(t)\,dt = 1\right\}.$$

We say that $\pi = \{t_0, t_1, \ldots, t_m\}$ is a *partition* of $[0, T]$, if $0 = t_0 < t_1 < \ldots < t_{m-1} < t_m = T$; we then write $|\pi| = \max_{1 \leqslant i \leqslant m} (t_i - t_{i-1})$. Now let $\pi_n = \{t_0{}^n, t_1{}^n, \ldots, t^n_{m(n)}\}$ ($n = 1, 2, \ldots$; the superscripts are indices) be a sequence of such partitions with $\lim_{n\to\infty} |\pi_n| = 0$, and write

$$A_n = \{F \in \Lambda_L : \sum_{i=1}^{m(n)} c(t_i{}^n) F(t_i{}^n)(t_i{}^n - t^n_{i-1}) = 1\}.$$

As we saw in Section 3, the appropriate member of $P_L$ for $A_n$ is

$$p_n(t) = \sum_{i=1}^{m(n)} c(t_i{}^n)(t_i{}^n - t^n_{i-1}) u(t_i{}^n - t).$$

4.1. Theorem. *$p_n$ converges uniformly to $p$ as $n \to \infty$.*

*Proof.* Reasoning by contradiction, suppose the uniform convergence does not occur. Then there is a $\delta > 0$, an increasing sequence, $\{n_j\}_{j=1}^{\infty}$, of positive integers, and a sequence $\{s_j\}_{j=1}^{\infty}$ of numbers in $[0, T]$, such that $|p_{n_j}(s_j) - p(s_j)| > \delta$. We can assume—taking a further subsequence, if necessary—that $s_j$ converges to some $s \in [0, T]$. We have

$$\begin{aligned} |p_{n_j}(s) - p(s)| &\geqslant |p_{n_j}(s_j) - p(s_j)| - |p_{n_j}(s) - p_{n_j}(s_j)| - |p(s_j) - p(s)| \\ &> \delta - |p_{n_j}(s) - p_{n_j}(s_j)| - |p(s_j) - p(s)|. \end{aligned}$$

Let us now prove the following.

Lemma. $\lim_{j\to\infty} |p_{n_j}(s) - p_{n_j}(s_j)| = 0.$

Fix $\epsilon > 0$. By integrability, there is a $K > 0$ such that, if

$$I = \{t \in [0,T]: |c(t)| > K\}, \quad \text{then} \quad \int_I |c(t)|\,dt < \epsilon/4.$$

With $K$ and $I$ thus defined, let

$$c_1(t) = \begin{cases} 0 & \text{when } t \in I \\ c(t) & \text{otherwise} \end{cases} \quad \text{and} \quad c_2(t) = \begin{cases} c(t) & \text{when } t \in I \\ 0 & \text{otherwise.} \end{cases}$$

Now, there exists a $J_1$ such that $j > J_1$ implies

$$\sum_{i=1}^{m(n)} |c_2(t_i^{nj})|(t_i^{nj} - t_{i-1}^{nj}) < 2 \int_0^T |c_2(t)|\,dt < \epsilon/2.$$

This is because the sum is a Riemann sum for the integral.

For each $j$, denote by $i(j)$ the smallest index, $\nu$, for which $\min(s_j, s) < t_\nu^{nj} \leqslant \max(s_j, s)$; and by $k(j)$, the largest. Then (with $h_i = t_i^{nj} - t_{i-1}^{nj}$),

$$\begin{aligned} |p_{n_j}(s) - p_{n_j}(s_j)| &\leqslant \sum_{i=i(j)}^{k(j)} |c(t_i^{nj})| h_i \\ &= \sum_{i=i(j)}^{k(j)} |c_1(t_i^{nj})|\, h_i + \sum_{i=i(j)}^{k(j)} |c_2(t_i^{nj})|\, h_i \\ &\leqslant K \sum_{i=i(j)}^{k(j)} h_i + \sum_{i=1}^{n} |c_2(t_i^{nj})|\, h_i \\ &\leqslant K(t_{i(j)}^{nj} - t_{i(j)-1}^{nj}) + K|s - s_j| + \tfrac{1}{2}\epsilon, \qquad \text{if } j > J_1. \end{aligned}$$

Since $s_j \to s$, there is a $J_2$, such that $j > J_2$ implies $|s - s_j| < \epsilon/(4K)$. Since $|\pi_n| \to 0$, there is a $J_3$, such that $|\pi_{n_j}| < \epsilon/(4K)$ whenever $j > J_3$. Thus, if $i > \max(J_1, J_2, J_3)$, then

$$|p_{n_j}(s) - p_{n_j}(s_j)| < [K\epsilon/(4K)] + [K\epsilon/(4K)] + \epsilon/2 = \epsilon.$$

This completes the proof of the lemma.

By the lemma, there is an $M_1$ such that $|p_{n_j}(s) - p_{n_j}(s_j)| < \delta/4$ whenever $j > M_1$. By the continuity of an integral with respect to the limits of integration, there is an $M_2$ for which $j > M_2$ implies $|p(s_j) - p(s)| < \delta/4$. Consequently, $|p_{n_j}(s) - p(s)| > \delta/2$ whenever $j > \max(M_1, M_2)$. But $p_{n_j}(s)$ $(j = 1, 2, \ldots)$ is a sequence of Riemann sums for $\int_0^T c_0(t)\,dt = p(s)$, where

$$c_0(t) = \begin{cases} 0 & \text{for } t < s \\ c(t) & \text{for } t \geqslant s. \end{cases}$$

Thus, we have a contradiction.

Q.E.D.

The last theorem has the following interpretation: If $c$ is Riemann integrable and $\{\pi_n\}_{n=1}^{\infty}$ is a sequence of partitions with $|\pi_n| \to 0$, and if $R_n(t)$ is the Riemann sum for $\int_t^T c(s)\,ds$, formed by restricting $\pi_n$ to $[t, T]$, then the convergence of $R_n(t)$ to $\int_t^T c(s)\,ds$ is uniform in $t$.

4.2. COROLLARY. *$A_n$ is nonvoid for all sufficiently large $n$.*

*Proof.* By assumption, $\int_0^T |p(t)|\,dt > 1/L$, and by 4.1,

$$\lim_{n\to\infty} \int_0^T |p_n(t)|\,dt = \int_0^T |p(t)|\,dt.$$

The conclusion then follows from 2.3 and from the results of Section 3.

Q.E.D.

In what follows, assume $\{\pi_n\}_{n=1}^{\infty}$ has been purged of any members for which $A_n$ is empty.

4.3. THEOREM. *If $\tilde{F}_n$ is the nucleus of $A_n$ $(n = 1, 2, \ldots)$, then $\tilde{F}_n$ converges uniformly to*

$$\tilde{F}(t) = L \int_0^t \tanh(\lambda p(s))\,ds,$$

*where $\lambda$ is chosen so that*

$$\int_0^T c(t)\,\tilde{F}(t)\,dt = 1.$$

*Proof.* The theorem follows from 4.1 and the representation

$$\tilde{F}_n(t) = L \int_0^t \tanh(\lambda_n p_n(s))\,ds,$$

where $\lambda_n$ is chosen so that

$$\sum_{i=1}^{m(n)} c(t_i^n)\,\tilde{F}_n(t_i^n)\,(t_i^n - t_{i-1}^n) = 1.$$

Q.E.D.

ACKNOWLEDGMENT

This text incorporates suggestions made by the referee.

REFERENCES

1. G. STRANG, The nucleus of a set. *Canadian Math. Bulletin* **11**, in press.
2. J. E. ISRAEL, Entropic determination of a metric nucleus. *J. Approx. Theory* **1** (1968), 189–198.