ZESZYTY NAUKOWE UNIWERSYTETU JAGIELLOŃSKIEGO
DCXXIII PRACE MATEMATYCZNE, ZESZYT 23 1982

On the convergence of the iterated Pilgerschritt transformation

by N. NETZER

The solution of the translation equation in a real Lie-group G can be defined as the construction of a continuous homomorphism $h\colon R \to G$ where $h(1) = f$ is a given element of G (cf. R. Liedl, [5]). In order to solve this problem in a more general case and if the restriction $h/[0, 1]$ is required to be homotopic to a given path $\varphi\colon [0, 1] \to G$ from the unit element to the element f, R. Liedl has proposed a method called the Pilgerschritt-transformation. In this paper we shall give sufficient conditions for this method to solve this problem.

Since a Lie-group G is locally isomorphic to a group of matrices and our conditions will be formulated using such a local isomorphism, we shall restrict the group G to a closed subgroup of the group $Gl(n, R)$ of the real $n \times n$-matrices. Let $f \in G$ be the given $n \times n$-matrix and let the C^1-path $\varphi\colon [0, 1] \to G$ with $\varphi(0) = E$ and $\varphi(1) = f$ represent the required homotopy class of the restriction $h/[0, 1]$ of the homomorphism $h\colon R \to G$ in question. Then it is possible to consider the solution of the matrix differential equation

$$\frac{\partial \varphi(t, \tau)}{\partial t} = \tau \varphi'(t) \varphi(t)^{-1} \varphi(t, \tau)$$

with the initial condition $\varphi(0, \tau) = E$ (where $\tau \in [0, 1]$ denotes a real parameter). The path $\varphi\colon [0, 1] \to G$ defined by $\varphi(\tau) = \varphi(1, \tau)$ is called the Pilgerschritt transform of φ (cf. R. Liedl [5]).

The iteration of the Pilgerschritt transformation gives rise to a sequence of paths $\varphi, \tilde{\varphi}, \tilde{\tilde{\varphi}}, \ldots, \overset{n}{\tilde{\varphi}}, \ldots$ Liedl's conjecture was that there exist weak conditions for the path φ implying the uniform convergence of this sequence to a path $\chi\colon [0, 1] \to G$, which is homotopic to φ and is the restriction of a continuous homomorphism $h\colon R \to G$.

In this paper we wish to show that there exists a positive real number L such that this conjecture is true, if $\|\varphi'(t)\| < L$ for each $t \in [0, 1]$ (e.g., $\|\cdot\|$ is the operator norm). If $\|\varphi'(t)\|$ is small enough for each $t \in [0, 1]$, by integration we see that $\|\varphi(t) - E\|$ is small. Thus $\varphi(t)$ can be written in the form

$$\varphi(t) = \exp(tD) \exp(A(t))$$

where $\exp(D) = f$ and $A(0) = A(1) = 0$.

Because of

$$\varphi'(t) = D\exp(tD)\exp(A(t))+\exp(tD)\frac{d}{dt}\exp(A(t))$$

we have $\|\varphi'(t)\|$ small iff $\|D\|$ and $\|A'(t)\|$ are small. Therefore it is sufficient to show the following

THEOREM: *For each $K\in(0,1)$ there exists a $\beta>0$ such that $\|D\|<\beta_1$ and $\max_{t\in[0,1]}\|A'(t)\|<\beta_1$ and $\beta_1<\beta$ implies that*

$$\varphi(t) = \exp(tD)\exp(A(t)),$$

where

$$\max_{t\in[0,1]}\|A'(t)\|<K\beta_1 .$$

We first prove the following

LEMMA: *Let $\delta>0$ be such that for $X, Y\in\mathscr{L}(G)$ (we identify the Lie-algebra of G with a subalgebra of the algebra of real $n\times n$-matrices) the Campbell–Baker–Hausdorff-series*

$$H(X, Y) = \log(\exp X\exp Y) = X+Y+\tfrac{1}{2}[X, Y]+\dots$$

converges whenever $\|X\|, \|Y\|<\delta$. We put

$$F(X, Y) = \frac{d}{dt}(H(tX, Y))|_{t=0} = X+\sum_{k\geq 1}\alpha_k[Y, \underbrace{[Y, \dots[Y, X]]\dots]}_{k \text{ brackets}}(F)$$

($\alpha_k = \dfrac{\xi_k}{k!}$ *where ξ_k is the k-th Bernoulli-number*) *and*

$$g(t, \tau) = \exp((1-\tau)tD)\left(\frac{d}{dt}\exp(A(t))\right)\exp(-A(t))\exp((\tau-1)tD).$$

If $B(t, \tau)$ is the solution of the matrix differential equation

$$(*)\qquad \frac{\partial Y(t, \tau)}{\partial t} = \tau F(g(t, \tau), Y(t, \tau)), \qquad Y(0, \tau) = 0 \quad \textit{for each } \tau\in[0, 1]$$

and if $\|g(t, \tau)\|<\delta$ and $\|B(t, \tau)\|<\delta$ for each $t, \tau\in[0, 1]$, then

$$\varphi(t, \tau) = \exp(\tau tD)\exp(B(t, \tau)).$$

Proof of the lemma: Differentiation shows that

$$\varphi'(t)\varphi(t)^{-1} = D+\exp(tD)\left(\frac{d}{dt}\exp(A(t))\right)\exp(-A(t))\exp(-tD).$$

On the other hand, setting

$$\psi(t) := \exp(\tau tD)\exp(B(t, \tau))$$

and

$$\Gamma := (\psi(t+h,\tau)-\psi(t,\tau))\psi(t,\tau)^{-1}$$

we have consecutively:

$$\Gamma = \exp(\tau(t+h)D)\exp\left(B(t,\tau)+h\frac{\partial}{\partial t}B(t,\tau)+o(h)\right)\exp(-B(t,\tau))\exp(-\tau tD)-E$$

by Taylor's formula,

$$\Gamma = \exp(\tau(t+h)D)\exp\big(B(t,\tau)+\tau hF(g(t,\tau),B(t,\tau))+o(h)\big)\times$$
$$\times\exp(-B(t,\tau))\exp(-\tau tD)-E$$

by substituting the differential equation,

$$\Gamma = \exp(\tau(t+h)D)\exp\big(H(\tau hg(t,\tau),B(t,\tau))+o(h)\big)\exp(-B(t,\tau))\exp(-\tau tD)-E$$

by using the definition of F,

$$\Gamma = \exp(\tau(t+h)D)\exp\big(H(H(\tau hg(t,\tau),B(t,\tau)),-B(t,\tau))+o(h)\big)\exp(-\tau tD)-E$$

by the Campbell–Baker–Hausdorff-formula,

$$\Gamma = \exp(\tau(t+h)D)\exp(\tau hg(t,\tau)+o(h))\exp(-\tau tD)-E$$
$$= \exp(\tau(t+h)D)(E+\tau hg(t,\tau)+o(h))\exp(-\tau tD)-E$$

by associativity of the matrix product, and

$$\Gamma = \exp(\tau hD)+\tau h\exp((t+\tau h)D)\left(\frac{d}{dt}\exp(A(t))\right)\exp(-A(t))\exp(-tD)+o(h)-E$$
$$= \tau h(D+\exp(tD)\left(\frac{d}{dt}\exp(A(t))\right)\exp(-A(t))\exp(-tD)+o(h)$$

by using the definition of g.

Therefore we have

$$\frac{\partial}{\partial t}\psi(t,\tau)\psi(t,\tau)^{-1} = \tau\varphi'(t)\varphi(t)^{-1}$$

and the lemma is proved.

Proof of the theorem. We use the lemma to calculate the Pilgerschritt-transform by solving the differential equation (*). In order to solve this differential equation we try to get $B(t,\tau) = \sum_{k\geq 1} B_k(t,\tau)$ which converges uniformly. For constructing the functions $B_k(t,\tau)$ $(k = 1, 2, \ldots)$ we consider the series $g(t,\tau) = \sum_{k\geq 1} G_k$, where

$$G_k = \sum_{n_1+n_2+n_3+n_4=k-1} a_{n_1n_2n_3n_4}(1-\tau)^{n_1+n_4}(tD)^{n_1}A(t)^{n_2}A'(t)A(t)^{n_3}(tD)^{n_4}$$

($a_{n_1n_2n_3n_4}$ are reals) is given by a straightforward calculation.

Now we substitute $\sum_{k\geq 1} B_k(t,\tau)$ for Y, $\sum_{k\geq 1} \frac{\partial B_k(t,\tau)}{\partial t}$ for $\frac{\partial Y}{\partial t}$ and (F) for $F(X, Y)$ in the differential equation. In order to interpret the right-hand side of the differential equation to be $\sum_{k\geq 1} \frac{\partial B_k(t,\tau)}{\partial t}$ it is necessary to consider this right-hand side as a series $\sum_{k\geq 1} L_k(t,\tau)$ in a suitable way.

Therefore we define

$$M_0 := \{(tD)^{n_1} A(t)^{n_2} A'(t) A(t)^{n_3} (tD)^{n_4} / n_1, n_2, n_3, n_4 \in N \cup \{0\}\},$$

$$M_0^* := \{\int_0^t Z(s)\,ds / Z \in M_0\}$$

and inductively

$$M_k := \{[X_1, [X_2, \ldots [X_n, Y]] \ldots] / n \in N, X_1, \ldots, X_n \in \bigcup_{l=0}^{k-1} M_l^*, Y \in M_0\},$$

$$M_k^* := \{\int_0^t Z(s)\,ds / Z \in M_k\},\ M := \bigcup_{k\geq 0} M_k \text{ and } M^* := \bigcup_{k\geq 0} M_k^*.$$

For $Z \in M$ we define $\deg Z$ by: $\deg Z := n_1+n_2+n_3+n_4+1$, for $Z \in M_0$;

$\deg Z := \sum_{j=1}^{n} \deg X_j + \deg Y$ for $Z \in M_k$ and

$$\deg \int_0^t \alpha Z(s)\,ds = \deg \alpha Z := \deg Z \ (\alpha \in R).$$

Further we define a Lie-bracket counting function ϱ by

$$\varrho(Z) := 0 \text{ if } Z \in M_0^* \cup M_0,\ \varrho(Z) := \varrho(X_1) + \ldots \varrho(X_n) + n \quad \text{for } Z \in M_k$$

and

$$\varrho\Big(\int_0^t \alpha Z(s)\,ds\Big) = \varrho(\alpha Z) := \varrho(Z) \quad \text{for } Z \in M \quad \text{and } \alpha \in R.$$

Now we are able to define $L_1(t,\tau) := \tau A'(t)$ and we set

$$B_1(t,\tau) := \int_0^t L_1(s,\tau)\,ds = \tau A(t),$$

because we want to have $\frac{\partial B_1(t,\tau)}{\partial t} = L_1(t,\tau)$. Inductively we define $L_k(t,\tau)$ to be the homogeneous part of degree k of the right-hand side. Therefore we have $L_k(t,\tau) = \tau(G_k(t))$ + the homogeneous part of degree k of the expression

$$\sum_{l=1}^{k-1} \alpha_l [\sum_{j=1}^{k-1} B_j(t,\tau), \ldots, [\sum_{j=1}^{k-1} B_j(t,\tau), \sum_{j=1}^{k-1} G_j(t)] \ldots]).$$

Hence we get $B_k(t,\tau) = \int_0^t L_k(s,\tau)\,ds$ and we have a formal expression for $B(t,\tau)$.

Convergence can be proved in the following way:
We can see

$$\frac{\partial B_k(t,\tau)}{\partial t} = \sum_{Z\in M,\deg Z=k} P_Z(\tau)Z$$

where $P_Z(\tau) = \sum_{i=1}^{n_Z} a_{iZ}\tau^i$ is a polynomial in τ with real coefficients. We define

$$P_Z^*(\tau) := \sum_{i=1}^{n_Z} |a_{iZ}|\tau^i .$$

In order to find the number $\beta>0$ of the theorem we calculate with an arbitrary $b>0$, supposing $\|D\|<b$ and $\max_{t\in[0,1]} \|A'(t)\|<b$. Then

$$\|\sum_{k\geqslant 1} B_k(t,\tau)\| \leqslant \sum_{k\geqslant 1} \|\sum_{Z\in M^*,\deg Z=k} P_Z(\tau)Z\| \leqslant$$

$$\leqslant \sum_{k\geqslant 1}\sum_{Z\in M^*,\deg Z=k} P_Z^*(\tau)\|Z\| \leqslant \sum_{k\geqslant 1}\sum_{Z\in M^*,\deg Z=k} P_Z^*(\tau)\frac{2^{\varrho(Z)}}{(\varrho(Z)+1)!}b^k$$

and

$$\left\|\sum_{k\geqslant 1}\frac{\partial B_k(t,\tau)}{\partial t}\right\| \leqslant \sum_{k\geqslant 1}\sum_{Z\in M,\deg Z=k} P_Z^*(\tau)\frac{2^{\varrho(Z)}}{\varrho(z)!}b^k .$$

First we wish to show that the series

$$R(b) = \sum_{k\geqq 1}\sum_{Z\in M,\deg Z=k} P_Z(1)\frac{2^{\varrho(Z)}}{\varrho(Z)!}b^k$$

and the series

$$S(b) = \sum_{k\geqq 1}\sum_{Z\in M,\deg Z=k} \frac{dP_Z(1)}{d\tau}\cdot\frac{2^{\varrho(Z)}}{\varrho(Z)!}b^k$$

have positive radii of convergence.

If b is less than each of these radii of convergence, the series $\sum_{k\geqq 1} B_k(t,\tau)$ and $\sum_{k\geqq 1}\left\|\frac{\partial B_k(t,\tau)}{\partial\tau}\right\|$ converge uniformly. Therefore we define

$$\gamma_k := \sum_{n_1+n_2+n_3+n_4=k-1} |a_{n_1n_2n_3n_4}|(1+\tau)^{n_1+n_4}b^k .$$

Using the identity

$$e^{(1-\tau)b}be^be^{-b}e^{(\tau-1)b} = \sum_{k\geqq 1}\sum_{n_1+n_2+n_3+n_4=k-1} a_{n_1n_2n_3n_4}(1-\tau)^{n_1+n_4}b^k$$

and $1+\tau\leqq 2$ we get a real number $\gamma = \sum_{k\geqq 1}\gamma_k \leqq be^{6b}$. Now consider the differential equation

$$v' = \tau(1+\sum_{k\geqq 1}|\alpha_k|2^kv^k)\gamma, \quad v(0)=0, \tag{**}$$

If we can show that the solution of (**) is a majorant of $R(b)$ for a b which is small enough, then we have ensured that $R(b)$ has a positive radius of convergence.

In order to show this we put $v = \sum_{k\geq 1} v_k$ and get

$$\sum_{k\geq 1} v_k' = \tau\Big(1+\sum_{k\geq 1} |\alpha_k| 2^k \big(\sum_{l\geq 1} v_l\big)^k\Big) \sum_{k\geq 1} \gamma_k .$$

A recursive calculation quite analogous to that of the functions $B_k(t, \tau)$ gives us $v_1' = \tau\gamma_1$, $v_1 = \tau t\gamma_1$ and
$v_k' = \tau\gamma_k^+$ the homogeneous part of degree k of

$$\tau\Big(\sum_{l=1}^{k-1} |\alpha_l| 2^l \big(\sum_{j=1}^{k-1} v_j\big)^l\Big) \sum_{l=1}^{k-1} \gamma_l$$

and $v_k = \int_0^t v_k'(s)\,ds$. The degree is given by interpreting the functions v_k and γ_k as polynomials in b.

Then we have

$$\sum_{Z\in M, \deg Z=k} P_Z^*(\tau) \frac{2^{\varrho(Z)}}{\varrho(Z)!} b^k \leq v_k' \quad \text{for } k = 1, 2, \ldots$$

The series $\sum_{k\geq 1} v_k$ and $\sum_{k\geq 1} v_k'$ converge uniformly. To see this, we solve the differential equation (**) in a different way. We put $w = \sum_{k\geq 1} w_k$ and have

$$\sum_{k\geq 1} w_k' = \tau\Big(1+\sum_{k\geq 1} |\alpha_k| 2^k \big(\sum_{l\geq 1} w_l\big)^k\Big)\gamma .$$

Again w_k is calculated recursively by $w_1' = \tau\gamma$, $w_1 = \tau t\gamma$, w_k' = homogeneous part of degree k of $\tau\big(1+\sum_{l=1}^{k-1} |\alpha_l| 2^l \big(\sum_{j=1}^{k-1} v_j\big)^l \gamma\big)$ and $w_k = \int_0^t w_k'(s)\,ds$. The degree is given by interpreting the functions w_k and w_k' as polynomials in γ.

Perron (cf. [2], [7]) has proved that the series $\sum_{k\geq 1} w_k$ and $\sum_{k\geq 1}^{k} w_k'$ converge uniformly and that $\sum_{k\geq 1} w_k$ is the solution of (**), if γ is small enough. The convergence of $R(b)$ is ensured, if b is small enough, because

$$\sum_{j=1}^{k} v_j(t) \leq \sum_{j=1}^{k} w_j(t) \quad \text{for } k = 1, 2, \ldots$$

$S(b)$ has also a positive radius of convergence because the polynomials P_Z^* have nonnegative coefficients. Therefore we have constructed the required majorants. Now choose a positive real β which is less than each of the radii of convergence of $R(b)$ and $S(b)$ such that $\beta e^{6\beta} < \delta$, $R(\beta) < \delta$ and

$$\sum_{k\geq 2} \sum_{Z\in M, \deg Z=k} \frac{dP_Z^*(1)}{dt} \frac{2^{\varrho(Z)}}{(\varrho(Z)+1)!} \beta^{k-1} < K .$$

Let be $0<\beta_1\leqq\beta$ and $\|D\|<\beta_1$ and $\max_{t\in[0,1]}\|A'(t)\|<\beta_1$. Then we have $\|g(t,\tau)\|<\delta$ and $\|B(t,\tau)\|<\delta$ and therefore the lemma may be used. For $B_1(1,\tau)=\tau A(1)=0$ we have

$$\left\|\frac{\partial}{\partial\tau}B(t,\tau)\right\|\leqq\sum_{k\geqq 2}\sum_{Z\in M,\deg Z=k}\frac{\partial}{\partial\tau}P_Z^*(\tau)\frac{2^{\varrho(Z)}}{(\varrho(Z)+1)!}\beta_1^k<K\beta_1\,.$$

Q.E.D.

Remarks:

1. If G is an abelian Lie group, then $\tilde{\varphi}=h/[0,1]$, cf. R. Liedl [5] (1).
2. If G is an abelian Banach–Lie group, then $\tilde{\varphi}=h/[0,1]$, cf. H. Reitberger [8] (1).
3. If G is a unipotent linear group, then there exists a natural number N such that $\overset{N}{\tilde{\varphi}}=h/[0,1]$, cf. K. Kuhnert [3] (1).
4. For solvable Lie groups K. Kuhnert [3] has proved that the sequence $\varphi,\tilde{\varphi},\tilde{\tilde{\varphi}},\ldots$ converges to $h/[0,1]$ under weaker conditions than those proposed in the theorem of this paper.
5. The following example shows that the condition $\|D\|<\beta$ in the theorem cannot be replaced by more special conditions on the path A (except the trivial condition $A(t)=0$ for each $t\in[0,1]$). Put $n=2$,

$$D=\begin{pmatrix}0&0\\0&c\end{pmatrix}\quad\text{and}\quad A(t)=\begin{pmatrix}0&a(t)\\0&0\end{pmatrix}.$$

Then

$$\tilde{A}(t)=\begin{pmatrix}0&\tilde{a}(t)\\0&0\end{pmatrix}.$$

$$\tilde{a}(t)=-t\int_0^1(t-1)ca(s)e^{(t-1)sc}ds\,.$$

Now choose reals α, β and c such that $0<\alpha<\beta<1$ and

$$K:=\alpha e^{(\beta-1)c\alpha}(e^{(\beta-1)c(\beta-\alpha)}-1)>1\,.$$

Further choose a: $[0,1]\to R$ such that $a(t)>\varepsilon>0$ for each $t\in[\alpha,\beta]$ and $a(t)\geqq 0$ for each $t\in[0,1]$. Then $\tilde{a}(t)\leqq 0$ for each $t\in[0,1]$ and

$$\begin{aligned}\tilde{a}(t)&\leqq -t\int_\alpha^\beta(t-1)ca(s)e^{(t-1)sc}ds\leqq-\varepsilon t\int_\alpha^\beta(t-1)ce^{(t-1)sc}ds\\&=-\varepsilon t(e^{(t-1)c\beta}-e^{(t-1)c\alpha})\leqq-\varepsilon\alpha^{(\beta-1)c\alpha}(e^{(\beta-1)c\alpha}(e^{(\beta-1)c(\beta-\alpha)}-1)<-K\varepsilon\end{aligned}$$

for $t\in[\alpha,\beta]$. An analogous calculation shows that

$$\tilde{\tilde{A}}(t)=\begin{pmatrix}0&\tilde{\tilde{a}}(t)\\0&0\end{pmatrix}$$

(1) φ is a path in G not necessarily having one of the properties required in the theorem or in the introduction to this paper.

where $\tilde{\tilde{a}}(t) \geqq 0$ for each $t \in [0, 1]$ and $\tilde{\tilde{a}}(t) > K^2 \varepsilon$ for $t \in [\alpha, \beta]$. This implies that $|\overset{r}{\tilde{a}}(t)| \geqq K^r \varepsilon$ for $t \in [\alpha, \beta]$ and the sequence $\varphi, \tilde{\varphi}, \tilde{\tilde{\varphi}}, \ldots$ diverges.

6. I have proved (unpublished) that $\overset{N+1}{\tilde{\varphi}} = \varphi$, if G is a nilpotent (not necessarily simply connected) Lie group whose degree of nilpotence is equal to N (1).

References

[1] N. Bourbaki, *Groups et algèbres de Lie*, Hermann, Paris (1971/72).

[2] E. Kamke, *Differentialgleichungen, Lösungsmethoden, Teubner*, Stuttgart (1977).

[3] K. Kuhnert, *Die Konvergenz des Pilgerschrittverfahrens für unipotente und auflösbare lineare Gruppen*, Berichte der mathematisch-statistischen Section im Forschungszentrum Graz, Ber. Nr. 87 (1978).

[4] R. Liedl, *Non-commutative calculus and Pilgerschritt transformation*, in this volume.

[5] R. Liedl, *Über eine Methode zur Lösung der Translationsgleichung*, Berichte der mathematisch-statischen Sektion im Forschungszentrum Graz, Ber. Nr. 84 (1978).

[6] N. Netzer, *Differentialgleichungen im Zusammenhang mit der Pilgerschrittransformation*, ibidem Ber. Nr. 86 (1978).

[7] O. Perron, *Sitzungsberichte der Heidelberger Akademie der Wissenschaften, mathematisch-naturwissenschaftliche Klasse Abt. A* (1919, 12., 2., 8. Abhandlung; 1920, 9. Abhandlung).

[8] H. Reitberger, *Pilgerschritttransformation und verallgemeinerte Liegruppen*, Berichte der mathematisch-statistischen Sektion im Forschungszentrum Graz, Ber. Nr. 85 (1978).

(1) φ is a parth in G not necessarily having one of the properties required in the theorem or in the introduction to this paper.