# MOTION OF A RIGID BODY UNDER RANDOM PERTURBATION 

MING LIAO ${ }^{1}$<br>Department of Mathematics, Auburn University, Auburn, AL 36849, USA. email: liaomin@auburn.edu<br>LONGMIN WANG<br>Department of Mathematics, Nankai University, Tianjin, China. email: wanglm@mail.nankai.edu.cn

Submitted 14 December 2004, accepted in final form 21 November 2005
AMS 2000 Subject classification: 58J65, 60H10
Keywords: rigid body, stochastic differential equation, rotation group, compact Lie group

## Abstract

We use stochastic analysis to study the random motion of a rigid body under a white noise perturbation. We obtain a formula for the angular velocity in an average sense and discuss the stability near a principle axis.

## 1 Introduction

In this paper, we study the random motion of a rigid body under a white noise perturbation. In an earlier paper [4] by the first author, the random motion of a rigid body is studied using a stochastic differential equation on the rotation group $S O(3)$ driven by a standard Brownian motion, which corresponds to an isotropic white noise perturbation. In the present paper, we will allow the driving Brownian motion to be non-standard, that is, there may be correlation among different components and the Brownian motion may be degenerate. This will allow us to consider the motion of a rigid body when the random perturbation is not uniform in different directions. In particular, the perturbation may be concentrated around a single axis. We will in fact work on a general compact Lie group $G$ without much additional effort, but will give interpretation only for $G=S O(3)$.
The rotation group $G=S O(3)$ is the group of $3 \times 3$ orthogonal matrices of determinant one. Its Lie algebra $\mathfrak{g}$, the tangent space $T_{e} G$ of $G$ at the identity element $e$, is the space $o(3)$ of $3 \times 3$ skew-symmetric matrices with Lie bracket $[X, Y]=X Y-Y X$. For $g \in G$, let $g X$ and $X g$ denote respectively the elements in $T_{g} G$ obtained from $X \in \mathfrak{g}$ by the left and the right translations. The adjoint action $\operatorname{Ad}$ of $G$ on its Lie algebra $\mathfrak{g}$ is defined by $\operatorname{Ad}(g) X=g X g^{-1}$. The canonical inner product on $\mathfrak{g}$, defined by $\langle X, Y\rangle=(1 / 2) \operatorname{Trace}\left(X Y^{\prime}\right)$, where $Y^{\prime}$ is the matrix transpose of $Y$, is $\operatorname{Ad}(G)$-invariant in the sense that $\langle\operatorname{Ad}(g) X, \operatorname{Ad}(g) Y\rangle=\langle X, Y\rangle$ for

[^0]any $g \in G$ and $X, Y \in \mathfrak{g}$. The space $o(3)$ may be identified with $\mathbb{R}^{3}$ with the inner products preserved such that any $X \in o(3)$ is a rotation in $\mathbb{R}^{3}$ in the sense that $e^{X}$ rotates $\mathbb{R}^{3}$ about the axis through $X$ (regarded as a vector in $\mathbb{R}^{3}$ ) by an angle $\|X\|=\langle X, X\rangle^{1 / 2}$.
The motion of a free rigid body fixed at the origin $o$ of $\mathbb{R}^{3}$ may be described by a smooth path $g(t)$ in $G=S O(3)$ with $g(0)=e$. The angular velocity $\boldsymbol{\omega}$ in the space and the angular velocity $\boldsymbol{\Omega}$ in the body, at time $t$ and with respect to $o$, are the elements of $\mathfrak{g}=o(3)$ (or vectors in $\mathbb{R}^{3}$ ) defined by $(d / d t) g(t)=\boldsymbol{\omega} g(t)=g(t) \boldsymbol{\Omega}$. They are respectively the angular velocities viewed by a person not moving with the body and by a person moving with the body. Note that they are related by $\boldsymbol{\omega}=\operatorname{Ad}(g(t)) \boldsymbol{\Omega}$.
The angular momenta $\mathbf{m}$ in the space and $\mathbf{M}$ in the body are elements in $\mathfrak{g}=o(3)$ (or vectors in $\mathbb{R}^{3}$ ) related by $\mathbf{m}=\operatorname{Ad}(g(t)) \mathbf{M}$. The inertia operator $L$ of the body respect to $o$ relates $\mathbf{M}$ to $\boldsymbol{\Omega}$ as $\mathbf{M}=L \boldsymbol{\Omega}$. This is a symmetric and positive definite linear operator: $\mathfrak{g} \rightarrow \mathfrak{g}$ in the sense that $\langle L X, Y\rangle=\langle X, L Y\rangle$ for $X, Y \in \mathfrak{g}$ and $\langle L X, X\rangle>0$ for nonzero $X$. Let $\Lambda=L^{-1}$. For a free rigid body, the angular momentum $\mathbf{m}$ is a constant. We have $(d / d t) g(t)=g(t) \boldsymbol{\Omega}=g(t) \Lambda(\mathbf{M})=g(t) \Lambda\left[\operatorname{Ad}\left(g(t)^{-1}\right) \mathbf{m}\right]$. Thus, the motion of a free rigid body may be determined by the following differential equation:
\[

$$
\begin{equation*}
d g(t)=g(t) \Lambda\left[\operatorname{Ad}\left(g(t)^{-1}\right) \mathbf{m} d t\right] \tag{1}
\end{equation*}
$$

\]

The reader is referred to [1] for more details on the dynamical theory of a rigid body.
To model a random perturbation that causes a continuous change in the position of the rigid body, and has independent and stationary effects over non-overlapping time intervals, we may replace $\mathbf{m} d t$ in (1) by $X_{0} d t+\sum_{i=1}^{3} E_{i} \circ d B_{t}^{i}$ for a 3-dimensional Brownian motion $B_{t}=$ $\left(B_{t}^{1}, B_{t}^{2}, B_{t}^{3}\right)$, where $\left\{E_{1}, E_{2}, E_{3}\right\}$ is an orthonormal basis of $\mathfrak{g}$ and od denotes the Stratonovich stochastic differential. As mentioned earlier, we will allow the driving Brownian motion $B_{t}$ to be non-standard, that is, its covariance matrix $\left\{a_{i j}\right\}$, with $a_{i j}=\operatorname{cov}\left(B_{1}^{i}, B_{1}^{j}\right)$, is not necessarily a diagonal matrix and may be degenerate. Thus, with $n=3$, the motion of the rigid body is a stochastic process $g_{t}$ in $G$ determined by the stochastic differential equation

$$
\begin{equation*}
d g_{t}=\sum_{i=1}^{n} g_{t} \Lambda\left[\operatorname{Ad}\left(g_{t}^{-1}\right) E_{i} \circ d B_{t}^{i}\right]+g_{t} \Lambda\left[\operatorname{Ad}\left(g_{t}^{-1}\right) X_{0} d t\right] \tag{2}
\end{equation*}
$$

with the initial condition $g_{0}=e$. The process $g_{t}$ is the motion of the rigid body with an initial angular momentum $X_{0}$ under the perturbation of the Brownian motion $B_{t}$. As for a free rigid body, we will call $X_{t}=\operatorname{Ad}\left(g_{t}^{-1}\right) X_{0}$ the angular momentum of the body at time $t$.
We will work under a more general framework. Let $G$ be an arbitrary compact connected Lie group with Lie algebra $\mathfrak{g}$. Fix an $\operatorname{Ad}(G)$-invariant inner product $\langle\cdot, \cdot\rangle$ on $\mathfrak{g}$ and let $\left\{E_{1}, E_{2}, \ldots, E_{n}\right\}$ be an orthonormal basis of $\mathfrak{g}$. Let $g_{t}$ be the process in $G$ determined by the stochastic differential equation (2) with the initial condition $g_{0}=e$, where $\Lambda: \mathfrak{g} \rightarrow \mathfrak{g}$ is a symmetric and positive definite linear map, and $B_{t}=\left(B_{t}^{1}, B_{t}^{2}, \ldots, B_{t}^{n}\right)$ is an $n$-dimensional Brownian motion with an arbitrary covariance matrix $\left\{a_{i j}\right\}$.
Since the path of the process $g_{t}$ is almost surely nowhere differentiable, the angular velocity cannot be defined in the usual sense, but may be defined in an average sense as follows. Write $g_{t+s}=g_{t} \exp \left(Z_{s}^{t}\right)$ for some $Z_{s}^{t} \in \mathfrak{g}$ with $Z_{0}^{t}=0$. Such a $Z_{s}^{t}$ always exists and is unique when $Z_{s}^{t}$ is sufficiently close to 0 , and in the case of $G=S O(3)$, it is the amount of rotation viewed in the body from time $t$ to time $t+s$. Let

$$
\begin{equation*}
L_{t}=\lim _{s \rightarrow 0} \frac{1}{s} E\left(Z_{s}^{t} \mid \mathcal{F}_{t}\right) \tag{3}
\end{equation*}
$$

where $\mathcal{F}_{t}$ is the $\sigma$-algebra generated by the Brownian motion up to time $t$, provided the limit in (3) exists almost surely. For $G=S O(3), E\left(Z_{s}^{t} \mid \mathcal{F}_{t}\right)$ is the average amount of rotation from $t$ to $t+s$, thus $L_{t}$ may be called the average angular velocity in the body at time $t$.
We will obtain an explicit formula for $L_{t}$ in Section 2. Unlike in the special case of isotropic perturbation in [4], $L_{t}$ may not vanish when $X_{0}=0$. In particular, we have the following interesting consequence: If the random perturbation is concentrated around a fixed axis, then the average random rotation will be around an axis that is orthogonal to the axis of perturbation. This result appears to be more significant than the one in [4] and the computation is also more complicated because the stochastic differential equation cannot be simplified due to the non-uniform perturbation. The reader is referred to a standard reference such as [3] for the stochastic calculus used in this paper.
In Section 3, we discuss the stability of a random rigid body, motivated by the stability of a free rigid body rotating near a principle axis, an axis along the direction of an eigenvector of $\Lambda$. When $X_{0}$ is along a principle axis and the perturbation is concentrate around $X_{0}$, the body will rotate randomly around $X_{0}$. For a general perturbation, because the stationary measure of the process $g_{t}$ can be shown to be the Haar measure on $G$, the body does not have a long run tendency to be rotating near a principle axis. However, the stability is reflected in the mean time it takes for the angular momentum $X_{t}=\operatorname{Ad}\left(g_{t}^{-1}\right) X_{0}$ to deviate a fixed distance from the principle axis. It will be shown that the mean time converges to infinity as the part of perturbation not around $X_{0}$ tends to zero and we will obtain a lower bound for the mean time which provides a quantitative control for this convergence. The proof is based on estimating the mean time when the energy is changed by $\delta>0$.

## 2 Average angular velocity

We will derive a stochastic differential equation satisfied by $Z_{s}^{t}$ as a process in time $s$. To simplify the notation, let

$$
\begin{equation*}
X_{t}=\operatorname{Ad}\left(g_{t}^{-1}\right) X_{0} \quad \text { and } \quad E_{t, i}=\operatorname{Ad}\left(g_{t}^{-1}\right) E_{i} \tag{4}
\end{equation*}
$$

For simplicity, $\Lambda(X)$ may be written as $\Lambda X$ for $X \in \mathfrak{g}$. Then (2) may be written as

$$
\begin{equation*}
d g_{t}=\sum_{i=1}^{n} g_{t} \Lambda E_{t, i} \circ d B_{t}^{i}+g_{t} \Lambda X_{t} d t \tag{5}
\end{equation*}
$$

Using the Stratonovich calculus, we may write $\left(\circ d g_{t}\right) g_{t}^{-1}+g_{t} \circ d g_{t}^{-1}=d\left(g_{t} g_{t}^{-1}\right)=0$, from which we obtain

$$
\begin{equation*}
d g_{t}^{-1}=-\sum_{i=1}^{n} \Lambda\left(E_{t, i}\right) g_{t}^{-1} \circ d B_{t}^{i}-\Lambda\left(X_{t}\right) g_{t}^{-1} d t \tag{6}
\end{equation*}
$$

and for $X \in \mathfrak{g}$,

$$
\begin{align*}
d \operatorname{Ad}\left(g_{t}^{-1}\right) X & =\left(\circ d g_{t}^{-1}\right) X g_{t}+g_{t}^{-1} X \circ d g_{t} \\
& =\sum_{i=1}^{n}\left[\operatorname{Ad}\left(g_{t}^{-1}\right) X, \Lambda E_{t, i}\right] \circ d B_{t}^{i}+\left[\operatorname{Ad}\left(g_{t}^{-1}\right) X, \Lambda X_{t}\right] d t \tag{7}
\end{align*}
$$

The differential of the exponential map at $X \in \mathfrak{g}$ is a linear map $D \exp (X): T_{X} \mathfrak{g} \rightarrow T_{\exp (X)} G$. By Theorem 1.7 in [2, chapter II], if $Y \in \mathfrak{g}$ is regarded as an element of $T_{X} \mathfrak{g}$, then

$$
\begin{equation*}
D \exp (X) Y=e^{X} \Gamma(X) Y \tag{8}
\end{equation*}
$$

where

$$
\begin{equation*}
\Gamma(X)=\frac{\operatorname{id}_{\mathfrak{g}}-e^{-\operatorname{ad}(X)}}{\operatorname{ad}(X)}=\operatorname{id}_{\mathfrak{g}}-\frac{1}{2!} \operatorname{ad}(X)+\frac{1}{3!} \operatorname{ad}(X)^{2}-\cdots \tag{9}
\end{equation*}
$$

and $\operatorname{ad}(X): \mathfrak{g} \rightarrow \mathfrak{g}$ is defined by $\operatorname{ad}(X) Y=[X, Y]$ (Lie bracket). The linear map $\Gamma(X): \mathfrak{g} \rightarrow \mathfrak{g}$ is invertible when $\|X\|$ is sufficiently small.
Fix $t \in \mathbb{R}_{+}$and apply the Stratonovich stochastic differentiation to $g_{t+s}=g_{t} \exp \left(Z_{s}^{t}\right)$ regarding both sides as processes in time $s$, we obtain $d g_{t+s}=g_{t} \exp \left(Z_{s}^{t}\right) \Gamma\left(Z_{s}^{t}\right) \circ d Z_{s}^{t}$. On the other hand, $d g_{t+s}=\sum_{i=1}^{n} g_{t+s} \Lambda E_{t+s, i} \circ d B_{t+s}^{i}+g_{t+s} \Lambda X_{t+s} d s$. This implies that

$$
\begin{align*}
& d Z_{s}^{t}=\sum_{i=1}^{n} \Gamma\left(Z_{s}^{t}\right)^{-1} \Lambda E_{t+s, i} \circ d B_{t+s}^{i}+\Gamma\left(Z_{s}^{t}\right)^{-1} \Lambda X_{t+s} d s \\
& =\sum_{i=1}^{n} \Gamma\left(Z_{s}^{t}\right)^{-1} \Lambda E_{t+s, i} d B_{t+s}^{i}+\Gamma\left(Z_{s}^{t}\right)^{-1} \Lambda X_{t+s} d s \\
& \quad+\frac{1}{2} \sum_{i=1}^{n} \Gamma\left(Z_{s}^{t}\right)^{-1} \Lambda\left(d E_{t+s, i}\right) \cdot d B_{t+s}^{i}+\frac{1}{2} \sum_{i=1}^{n} D \Gamma\left(Z_{s}^{t}\right)^{-1}\left(d Z_{s}^{t}, \Lambda E_{t+s, i}\right) \cdot d B_{t+s}^{i} \tag{10}
\end{align*}
$$

where $d B_{t+s}^{i}$ without the circle o is the Itô stochastic differential as usual, and for any two continuous semi-martingales $X_{t}$ and $Y_{t}, d X_{t} \cdot d Y_{t}$ denotes the differential of the associated quadratic covariance process. Moreover, if $V$ is a finite dimensional linear space, $L(V)$ is the space of linear endomorphisms on $V$ and $F: V \rightarrow L(V)$ is a smooth map, then for $x, y, z \in V$, $D F(x)(z, y)$ denotes the derivative $\left.(d / d t) F(x+t z) y\right|_{t=0}$. We obtain, by (7),

$$
\begin{align*}
d Z_{s}^{t}= & \sum_{i=1}^{n} \Gamma\left(Z_{s}^{t}\right)^{-1} \Lambda E_{t+s, i} d B_{t+s}^{i}+\Gamma\left(Z_{s}^{t}\right)^{-1} \Lambda X_{t+s} d s+\frac{1}{2} \sum_{i, j=1}^{n} \Gamma\left(Z_{s}^{t}\right)^{-1} \Lambda\left[E_{t+s, i}, \Lambda E_{t+s, j}\right] a_{i j} d s \\
& +\frac{1}{2} \sum_{i, j=1}^{n} D \Gamma\left(Z_{s}^{t}\right)^{-1}\left(\Gamma\left(Z_{s}^{t}\right)^{-1} \Lambda E_{t+s, i}, \Lambda E_{t+s, j}\right) a_{i j} d s \tag{11}
\end{align*}
$$

Because $Z_{0}^{t}=0, Z_{s}^{t}$ is equal to the integral $\int_{0}^{s}$ of the right hand side of (11).
Note that $Z_{s}^{t}=\eta\left(g_{t}^{-1} g_{t+s}\right)$ for a Borel measurable function $\eta: G \rightarrow \mathfrak{g}$ with $\exp \circ \eta=\mathrm{id}_{G}$ that is smooth only in a neighborhood $U$ of $e$. Therefore, $Z_{s}^{t}$ may not be a semi-martingale for all time $s$ and the usual rule of the stochastic calculus that has been used here may not be applied. Moreover, $\Gamma(Z)$ is invertible only when $\|Z\|$ is sufficiently small and hence $\Gamma\left(Z_{s}^{t}\right)^{-1}$ in (11) may not be meaningful. However, there is a constant $c>0$ such that if $\|Z\|<c$, then $\Gamma(Z)$ is invertible and $e^{Z} \in U$. It follows that $Z_{s}^{t}$ is a semi-martingale and $\Gamma\left(Z_{s}^{t}\right)$ is invertible for $s<\tau$, where $\tau$ is the first time $s$ when $\left\|Z_{s}^{t}\right\| \geq c$. We see that (11) is valid for $s<\tau$.
We may modify the $\Gamma(Z)$ for $\|Z\| \geq c$ so that $\Gamma(Z)$ is an invertible linear operator for all $Z \in \mathfrak{g}$, both $\Gamma(Z)$ and $\Gamma(Z)^{-1}$ are smooth in $Z$, and $\Gamma(Z)=\mathrm{id}_{\mathfrak{g}}$ when $\|Z\|$ is sufficiently large. Then the operator norms of $\Gamma(Z)$ and $\Gamma(Z)^{-1}$, and their derivatives, are bounded in $Z$. With $\Gamma$ thus modified, all the integrands in the stochastic differential equation (11) are bounded and hence we can solve it to obtain a process $\tilde{Z}_{s}^{t}$ in $\mathfrak{g}$ with $\tilde{Z}_{0}^{t}=0$. Then $Z_{s}^{t}=\tilde{Z}_{s}^{t}$ for $s<\tau$. We claim that

$$
\begin{equation*}
\lim _{s \rightarrow 0} \frac{1}{s} E\left(Z_{s}^{t} \mid \mathcal{F}_{t}\right)=\lim _{s \rightarrow 0} \frac{1}{s} E\left(\tilde{Z}_{s}^{t} \mid \mathcal{F}_{t}\right) \tag{12}
\end{equation*}
$$

provided the limit on the right hand side exists. To prove (12), noting that $Z_{s}^{t}$ is bounded but
$\tilde{Z}_{s}^{t}$ may not be so, we need to prove

$$
\lim _{s \rightarrow 0} \frac{1}{s} P(s \geq \tau)=\lim _{s \rightarrow 0} \frac{1}{s} P\left(\sup _{0 \leq u \leq s}\left\|\tilde{Z}_{u}^{t}\right\| \geq c\right)=0 \quad \text { and } \quad \lim _{s \rightarrow 0} \frac{1}{s} E\left[\left\|\tilde{Z}_{s}^{t}\right\| ; s \geq \tau\right]=0
$$

where $E[Z ; B]=\int_{B} Z d P$. We will only prove the second equation here because the proof of the first is similar. Since all the integrands on the right hand side of (11), with $Z_{s}^{t}$ replaced by $\tilde{Z}_{s}^{t}$, are bounded, and except the first term, the integrals $\int_{0}^{s}$ of the three other terms tend to 0 as $s \rightarrow 0$, hence, it suffices to show $\frac{1}{s} E\left(M_{s}^{*} ; M_{s}^{*}>c\right) \rightarrow 0$ as $s \rightarrow 0$ for any constant $c>0$, where $M_{s}=\int_{0}^{s} H_{u} d B_{u}, H_{u}$ is a bounded continuous adapted process, $B_{u}$ is a standard Brownian motion, and $M_{s}^{*}=\sup _{0 \leq u \leq s}\left|M_{u}\right|$. By a well known moment inequality for martingales, see III.Theorem 3.1 in $[3], E\left[\left(M_{s}^{*}\right)^{4}\right] \leq b E\left[\langle M, M\rangle_{s}^{2}\right]=b E\left[\left(\int_{0}^{s} H_{u}^{2} d u\right)^{2}\right]$ for some constant $b>0$. Then

$$
\frac{1}{s} E\left[M_{s}^{*} ; M_{s}^{*}>c\right] \leq \frac{1}{c^{3} s} E\left[\left(M_{s}^{*}\right)^{4}\right] \leq \frac{b}{c^{3} s} E\left[\left(\int_{0}^{s} H_{u}^{2} d u\right)^{2}\right] \rightarrow 0 \text { as } s \rightarrow 0
$$

This proves (12). Therefore, to determine the limit $L_{t}$ in (3), we may replace $Z_{s}^{t}$ by $\tilde{Z}_{s}^{t}$ which satisfies (11) for all $s$. For simplicity, we will write $Z_{s}^{t}$ for $Z_{s}^{t}$ and we may assume (11) holds for all $s$ in the following.
By (9),

$$
\begin{aligned}
\Gamma(Z+Y) X & =X-(1 / 2!)[Z+Y, X]+(1 / 3!)[Z+Y,[Z+Y, X]]+\ldots \\
& =\Gamma(Z) X+\Gamma(Y) X+O(\|Y\| \cdot\|Z\|)
\end{aligned}
$$

It follows that

$$
\begin{equation*}
D \Gamma(Z)(Y, X)=\left.\frac{d}{d t} \Gamma(Z+t Y) X\right|_{t=0}=-\frac{1}{2}[Y, X]+O(\|Z\|) \tag{13}
\end{equation*}
$$

In particular, $D \Gamma(Z)(X, X)=O(\|Z\|)$. Differentiating the identity $\Gamma(Z) \Gamma(Z)^{-1} X=X$ with respect to $Z$ yields $D \Gamma(Z)\left(Y, \Gamma(Z)^{-1} X\right)+\Gamma(Z) D \Gamma(Z)^{-1}(Y, X)=0$ and hence

$$
\begin{equation*}
D \Gamma(Z)^{-1}(Y, X)=-\Gamma(Z)^{-1} D \Gamma(Z)\left(Y, \Gamma(Z)^{-1} X\right) \tag{14}
\end{equation*}
$$

Then

$$
\begin{aligned}
& \sum_{i, j=1}^{n} D \Gamma\left(Z_{u}^{t}\right)^{-1}\left(\Gamma\left(Z_{u}^{t}\right)^{-1} \Lambda E_{t+u, i}, \Lambda E_{t+u, j}\right) a_{i j} \\
= & -\sum_{i, j=1}^{n} \Gamma\left(Z_{u}^{t}\right)^{-1} D \Gamma\left(Z_{u}^{t}\right)\left(\Gamma\left(Z_{u}^{t}\right)^{-1} \Lambda E_{t+u, i}, \Gamma\left(Z_{u}^{t}\right)^{-1} \Lambda E_{t+u, j}\right) a_{i j} \\
= & \frac{1}{2} \sum_{i, j=1}^{n} \Gamma\left(Z_{u}^{t}\right)^{-1}\left\{\left[\Gamma\left(Z_{u}^{t}\right)^{-1} \Lambda E_{t+u, i}, \Gamma\left(Z_{u}^{t}\right)^{-1} \Lambda E_{t+u, j}\right]+O\left(\left\|Z_{u}^{t}\right\|\right)\right\} a_{i j} \\
= & \Gamma\left(Z_{u}^{t}\right)^{-1} O\left(\left\|Z_{u}^{t}\right\|\right)=O\left(\left\|Z_{u}^{t}\right\|\right)
\end{aligned}
$$

Therefore, the integral of the last term on the right hand side of (11) is equal to $\int_{0}^{s} O\left(\left\|Z_{u}^{t}\right\|\right) d u$. The integral of the first term is an Itô integral, and because its integrand is bounded, it vanishes after taking the expectation. It follows that

$$
\begin{gathered}
\frac{1}{s} E\left(Z_{s}^{t} \mid \mathcal{F}_{t}\right)=E\left[\left.\frac{1}{s} \int_{0}^{s} \Gamma\left(Z_{u}^{t}\right)^{-1} \Lambda X_{t+u} d u \right\rvert\, \mathcal{F}_{t}\right] \\
+E\left\{\left.\frac{1}{s} \int_{0}^{s} \frac{1}{2} \sum_{i, j=1}^{n} \Gamma\left(Z_{u}^{t}\right)^{-1} \Lambda\left[E_{t+u, i}, \Lambda E_{t+u, j}\right] a_{i j} d u \right\rvert\, \mathcal{F}_{t}\right\}+E\left[\left.\frac{1}{s} \int_{0}^{s} O\left(\left\|Z_{u}^{t}\right\|\right) d u \right\rvert\, \mathcal{F}_{t}\right]
\end{gathered}
$$

which converges to $\Lambda X_{t}+\frac{1}{2} \sum_{i, j=1}^{n} \Lambda\left[E_{t, i}, \Lambda E_{t, j}\right] a_{i j}$ as $s \rightarrow 0$. The following result is proved.
Theorem 1 The average angular velocity in the body at time $t, L_{t}=\lim _{s \rightarrow 0} E\left(Z_{s}^{t} \mid \mathcal{F}_{t}\right)$, exists almost surely and is given by

$$
\begin{equation*}
L_{t}=\Lambda \operatorname{Ad}\left(g_{t}^{-1}\right) X_{0}+\frac{1}{2} \sum_{i, j=1}^{n} a_{i j} \Lambda\left[\operatorname{Ad}\left(g_{t}^{-1}\right) E_{i}, \Lambda \operatorname{Ad}\left(g_{t}^{-1}\right) E_{j}\right] \tag{15}
\end{equation*}
$$

Let $L_{t}^{B}$ be the second term on the right hand side of (15), which may be regarded as the contribution to the average angular velocity $L_{t}$ from the Brownian motion $B_{t}$, whereas the first term is the contribution from the initial angular momentum $X_{0}$. As mentioned earlier, by the result in [4], $L_{t}^{B}=0$ if $a_{i j}=\varepsilon \delta_{i j}$ for some $\varepsilon>0$. We can verify this directly as follows. Because $\operatorname{Ad}(g)$ is an orthogonal transformation on $\mathfrak{g}$ and the matrix $\left\{\Lambda_{i j}\right\}$ representing $\Lambda$ is symmetric,

$$
L_{t}^{B}=\frac{\varepsilon}{2} \sum_{i=1}^{n} \Lambda\left[\operatorname{Ad}\left(g_{t}^{-1}\right) E_{i}, \Lambda \operatorname{Ad}\left(g_{t}^{-1}\right) E_{i}\right]=\frac{\varepsilon}{2} \sum_{i=1}^{n} \Lambda\left[E_{i}, \Lambda E_{i}\right]=\frac{\varepsilon}{2} \Lambda \sum_{i, j=1}^{n} \Lambda_{i j}\left[E_{i}, E_{j}\right]=0
$$

Thus, $L_{t}^{B}=0$.
Recall that the covariance matrix $\left\{a_{i j}\right\}$ is allowed to be degenerate. We now consider the case when $\left\{a_{i j}\right\}$ has only one nonzero eigenvalue $\varepsilon>0$ (counting multiplicity). By choosing the orthonormal basis $\left\{E_{1}, \ldots, E_{n}\right\}$ of $\mathfrak{g}$ properly, we may assume: $a_{11}=\varepsilon$ and all other $a_{i j}=0$. Then $L_{t}^{B}=(1 / 2) \Lambda\left[\operatorname{Ad}\left(g_{t}^{-1}\right) E_{1}, \Lambda \operatorname{Ad}\left(g_{t}^{-1}\right) E_{1}\right]$. Because $\langle\Lambda[X, \Lambda X], X\rangle=\langle[X, \Lambda X], \Lambda X\rangle=$ $-\langle X,[\Lambda X, \Lambda X]\rangle=0$,

$$
\begin{equation*}
\left\langle L_{t}^{B}, \operatorname{Ad}\left(g_{t}^{-1}\right) E_{1}\right\rangle=0 \tag{16}
\end{equation*}
$$

that is, $L_{t}^{B}$ is orthogonal to $\operatorname{Ad}\left(g_{t}^{-1}\right) E_{1}$. Note that $L_{t}^{B}=0$ implies $\left[E_{1}, \Lambda E_{1}\right]=0$. On the other hand, if $\left[E_{1}, \Lambda E_{1}\right]=0$, then $g_{t}=\exp \left(t \Lambda E_{1}\right)$ is a solution of (2) in the present case and $\operatorname{Ad}\left(g_{t}^{-1}\right) E_{1}=E_{1}$, which implies $L_{t}^{B}=0$. To summarize, we have the following result.

Theorem 2 Assume $a_{11}>0$ and all other $a_{i j}=0$ (this means that the perturbation is concentrated around $E_{1}$ ). Then $L_{t}^{B}$ is orthogonal to $\operatorname{Ad}\left(g_{t}^{-1}\right) E_{1}$. Moreover, $L_{t}^{B}=0$ if and only if $\left[E_{1}, \Lambda E_{1}\right]=0$.

Note that for $G=S O(3),\left[E_{1}, \Lambda E_{1}\right]=0$ if and only if $\Lambda E_{1}$ is proportional to $E_{1}$, therefore, $L_{t}^{B}=0$ if and only if $E_{1}$ is an eigenvector of $\Lambda$.

## 3 Stability

By (2), $g_{t}$ is a diffusion process in $G$ with generator given by $L=(1 / 2) \sum_{i, j=1}^{n} a_{i j} U_{i} U_{i}+U_{0}$, where $U_{i}(g)=g\left[\Lambda \operatorname{Ad}\left(g^{-1}\right) E_{i}\right]$ and $U_{0}=g\left[\Lambda \operatorname{Ad}\left(g^{-1}\right) X_{0}\right]$ for $g \in G$. Let $P_{t}$ be its transition semigroup. A probability measure $\mu$ on $G$ is called a stationary measure of the diffusion process $g_{t}$ if $\int_{G} \mu(d x) P_{t}(x, \cdot)=\mu$ for any $t>0$. Thus, if the process $g_{t}$ is started with a stationary measure as the initial distribution, then it will have the same distribution at all time $t$. It is easy to show that the normalized Haar measure $\rho$ on $G$ is a stationary measure if and only if $L^{*} 1=0$, where $L^{*}$ is the operator dual to $L$ under $\rho$, that is, $\int(L \phi) \psi d \rho=\int \phi\left(L^{*} \psi\right) d \rho$ for $\phi, \psi \in C^{\infty}(G)$. The proof of Proposition 2 in [4] shows that $U_{i}^{*}=-U_{i}$ for $0 \leq i \leq n$. It follows that $L^{*} 1=0$ and hence $\rho$ is a stationary measure of $g_{t}$.

The perturbation is called non-degenerate if the matrix $\left\{a_{i j}\right\}$ is so. In this case, $\rho$ is the unique stationary measure of $g_{t}$ and hence, by the compactness of $G$, the distribution of $g_{t}$ converges to $\rho$ weakly as $t \rightarrow \infty$. For a degenerate perturbation, there may be other stationary measures and the distribution of $g_{t}$ may not converge to $\rho$.
As an example, assume the perturbation is concentrated around $X_{0}$. Without loss of generality, we may assume $X_{0}=b E_{1}$ for some constant $b, a_{11}>0$ and all other $a_{i j}=0$. The equation (2) now becomes $d g_{t}=g_{t} \Lambda \operatorname{Ad}\left(g_{t}^{-1}\right) E_{1} \circ d\left(B_{t}^{1}+b t\right)$. It is easy to show by stochastic calculus that the process $X_{t}=\operatorname{Ad}\left(g_{t}^{-1}\right) X_{0}$ in $\mathfrak{g}$ preserves the energy function

$$
\begin{equation*}
T(X)=\langle X, \Lambda X\rangle \tag{17}
\end{equation*}
$$

Moreover, by the Itô formula, the process $g_{t}$ is supported by the path of a free rigid body in the sense that $g_{t}=\Phi\left(B_{t}^{1}+b t\right)$, where $\Phi(t)$ is the solution of the ordinary differential equation (1) with $\mathbf{m}=E_{1}$ and $g(0)=e$. In the case of $G=S O(3)$, the angular momentum $\operatorname{Ad}\left(\Phi(t)^{-1}\right) X_{0}$ traces out a closed curve in $\mathfrak{g}=o(3) \equiv \mathbb{R}^{3}$ and $\Phi(t)$ is supported by a two-dimensional torus in $G=S O(3)$ (a sub-manifold, but not a subgroup), see the discussion in Section 29 of [1] for more details. When $X_{0}$ is an eigenvector of $\Lambda$, the body rotates randomly around $X_{0}$ and the torus reduces to a simple closed curve in $G=S O(3)$.
The rotating axis or the angular momentum of a free rigid body is stable near a principle axis associated to the largest or the smallest eigenvalue of $\Lambda$ in the sense that if it starts near such an axis, it will remain close forever. For a random rigid body, the angular momentum $X_{t}=\operatorname{Ad}\left(g_{t}^{-1}\right) X_{0}$ is a process in the sphere $S_{M}$ of radius $M=\left\|X_{0}\right\|$ centered at the origin in $\mathfrak{g}$. If the perturbation is non-degenerate, then $g_{t}$ converges in distribution to $\rho$ and hence $X_{t}$ converges in distribution to the uniform distribution on $S_{M}$. This means that $X_{t}$ does not have a long time tendency to be near a principle axis. However, assuming $X_{0}$ is along a principle axis, the stability around this axis is reflected in the mean time $E\left(\tau_{\delta}\right)$ when $X_{t}$ deviates from $X_{0}$ by a fixed distance. By the discussion in the previous paragraph, if the perturbation is concentrated around $X_{0}$, then $X_{t}=X_{0}$ for all $t>0$ and hence $E\left(\tau_{\delta}\right)=\infty$. We will show that in general $E\left(\tau_{\delta}\right) \rightarrow \infty$ when the part of the perturbation not around $X_{0}$ tends to zero and we will obtain a lower bound for $E\left(\tau_{\delta}\right)$ which provides a quantitative control for this convergence. Assume $X_{0} \neq 0$ is an eigenvector of $\Lambda$ associated to either the largest or the smallest eigenvalue. Without loss of generality, we may assume $X_{0}=M E_{1}$ with $M=\left\|X_{0}\right\|$. By (7), the symmetry of $\Lambda$ and the orthogonality of $\left[X_{t}, \Lambda X_{t}\right]=M\left[X_{t}, \Lambda E_{t, 1}\right]$ to $\Lambda X_{t}$,

$$
\begin{align*}
d T\left(X_{t}\right) & =2 \sum_{i=2}^{n}\left\langle\left[X_{t}, \Lambda E_{t, i}\right], \Lambda X_{t}\right\rangle \circ d B_{t}^{i}=2 M^{2} \sum_{i=2}^{n}\left\langle\left[E_{t, 1}, \Lambda E_{t, i}\right], \Lambda E_{t, 1}\right\rangle \circ d B_{t}^{i} \\
& =2 M^{2} \sum_{i=2}^{n}\left\langle\left[E_{t, 1}, \Lambda E_{t, i}\right], \Lambda E_{t, 1}\right\rangle d B_{t}^{i}+M^{2} \sum_{i=2}^{n} \sum_{j=1}^{n} a_{i j} F_{i j}\left(g_{t}\right) d t \tag{18}
\end{align*}
$$

where

$$
\begin{align*}
F_{i j}(g)= & \left\langle\left[\left[\operatorname{Ad}\left(g^{-1}\right) E_{1}, \Lambda \operatorname{Ad}\left(g^{-1}\right) E_{j}\right], \Lambda \operatorname{Ad}\left(g^{-1}\right) E_{i}\right], \Lambda \operatorname{Ad}\left(g^{-1}\right) E_{1}\right\rangle \\
& +\left\langle\left[\operatorname{Ad}\left(g^{-1}\right) E_{1}, \Lambda\left[\operatorname{Ad}\left(g^{-1}\right) E_{i}, \Lambda \operatorname{Ad}\left(g^{-1}\right) E_{j}\right]\right], \Lambda \operatorname{Ad}\left(g^{-1}\right) E_{1}\right\rangle \\
& +\left\langle\left[\operatorname{Ad}\left(g^{-1}\right) E_{1}, \Lambda \operatorname{Ad}\left(g^{-1}\right) E_{i}\right], \Lambda\left[\operatorname{Ad}\left(g^{-1}\right) E_{1}, \Lambda \operatorname{Ad}\left(g^{-1}\right) E_{j}\right]\right\rangle . \tag{19}
\end{align*}
$$

For $\delta>0$, let $\tau_{\delta}$ be the first time the change in $X_{t}$ causes an energy change of $\delta$, that is,

$$
\begin{equation*}
\tau_{\delta}=\inf \left\{t>0 ; \quad\left|T\left(X_{0}\right)-T\left(X_{t}\right)\right|=\delta\right\} \tag{20}
\end{equation*}
$$

By the usual convention, the inf of an empty set is defined to be $\infty$.
Since first term on the right hand side of (18) is the Itô stochastic differential of a martingale, if $E\left(\tau_{\delta}\right)<\infty$, then

$$
\begin{equation*}
\pm \delta=E\left[T\left(X_{0}\right)-T\left(X_{\tau_{\delta}}\right)\right]=M^{2} E\left[\int_{0}^{\tau_{\delta}} \sum_{i=2}^{n} \sum_{j=1}^{n} a_{i j} F_{i j}\left(g_{t}\right) d t\right] \tag{21}
\end{equation*}
$$

Recall $M=\left\|X_{0}\right\|$. Assume $M>0$. Let $C=\sup \left\{\left|F_{i j}(g)\right| ; g \in G\right.$ and $\left.1 \leq i, j \leq n\right\}$. By (21), $\delta \leq M^{2} C E\left(\tau_{\delta}\right) \sum_{i=2}^{n} \sum_{j=1}^{n}\left|a_{i j}\right|$. It follows that

$$
\begin{equation*}
E\left(\tau_{\delta}\right) \geq \frac{\delta}{M^{2} C\left[\sum_{i=2}^{n} \sum_{j=1}^{n}\left|a_{i j}\right|\right]} \tag{22}
\end{equation*}
$$

By (22) and noting $\left|a_{i j}\right| \leq \sqrt{a_{i i}} \sqrt{a_{j j}}$, we obtain the following result.
Theorem 3 Assume $X_{0} \neq 0$ is an eigenvector of $\Lambda$ associated to either the largest or the smallest eigenvalue. Let $\tau_{\delta}$ be the first time when the energy changes by $\delta>0$, defined by (20). Then $E\left(\tau_{\delta}\right) \rightarrow \infty$ as $a_{i i} \rightarrow 0$ for $2 \leq i \leq n$ and $a_{11}$ remains bounded.

Note that if in Theorem 3 the extreme eigenvalue is simple, then for any neighborhood $U$ of $X_{0}$ in $\mathfrak{g}$, the first exit time $\tau_{U}$ of $X_{t}$ from $U$ is larger than $\tau_{\delta}$ for a sufficiently small $\delta>0$. Consequently, $E\left(\tau_{U}\right) \rightarrow \infty$ as $a_{i i} \rightarrow 0$ for $2 \leq i \leq n$ and $a_{11}$ remains bounded.
We may replace $C$ in (22) by a larger constant for which an explicit expression may be obtained. As before, we will assume the orthonormal basis $\left\{E_{1}, \ldots, E_{n}\right\}$ of $\mathfrak{g}$ is chosen so that $X_{0}=M E_{1}$ is an eigenvector of $\Lambda$ associated to either the largest or the smallest eigenvalue $\lambda_{1}$. We may assume $E_{2}, \ldots, E_{n}$ are also eigenvectors of $\Lambda$ associated to eigenvalues $\lambda_{2}, \ldots, \lambda_{n}$. Let $C_{i j}^{k}$ be the structure constants of the Lie group $G$ under this basis given by

$$
\begin{equation*}
\left[E_{i}, E_{j}\right]=\sum_{k=1}^{n} C_{i j}^{k} E_{k} \tag{23}
\end{equation*}
$$

We have $C_{i j}^{k}=-C_{j i}^{k}$, and because the basis is orthonormal, $C_{i j}^{k}=-C_{i k}^{j}$. Let $\left\{b_{i j}(g)\right\}$ be the orthogonal matrix defined by $\operatorname{Ad}\left(g^{-1}\right) E_{i}=\sum_{j=1}^{n} b_{i j}(g) E_{j}$. By (19),

$$
\begin{align*}
F_{i j}= & \sum_{p, q, r, s=1}^{n} b_{i p} b_{j q} b_{1 r} b_{1 s}\left\{\left\langle\left[\left[\left[E_{r}, \Lambda E_{q}\right], \Lambda E_{p}\right], \Lambda E_{s}\right]\right\rangle+\left\langle\left[E_{r}, \Lambda\left[E_{p}, \Lambda E_{q}\right]\right], \Lambda E_{s}\right\rangle\right. \\
& \left.\quad+\left\langle\left[E_{r}, \Lambda E_{p}\right], \Lambda\left[E_{s}, \Lambda E_{q}\right]\right\rangle\right\} \\
= & \sum_{p, q, r, s} \sum_{k} b_{i p} b_{j q} b_{1 r} b_{1 s}\left\{C_{r q}^{k} C_{k p}^{s} \lambda_{p} \lambda_{q} \lambda_{s}+C_{p q}^{k} C_{r k}^{s} \lambda_{q} \lambda_{k} \lambda_{s}+C_{r p}^{k} C_{s q}^{k} \lambda_{p} \lambda_{q} \lambda_{k}\right\} \tag{24}
\end{align*}
$$

By the Schwartz inequality and the fact that $\sum_{j} b_{i j}^{2}=1$,

$$
\begin{align*}
\left|F_{i j}\right| & \leq(3 n)^{1 / 2}\left\{\sum_{p, q, r, s, k=1}^{n}\left[\left(C_{r q}^{k} C_{k p}^{s}\right)^{2} \lambda_{p}^{2} \lambda_{q}^{2} \lambda_{s}^{2}+\left(C_{p q}^{k} C_{r k}^{s}\right)^{2} \lambda_{q}^{2} \lambda_{k}^{2} \lambda_{s}^{2}+\left(C_{r p}^{k} C_{s q}^{k}\right)^{2} \lambda_{p}^{2} \lambda_{q}^{2} \lambda_{k}^{2}\right]\right\}^{1 / 2} \\
& =\sqrt{3 n \sum_{p, q, r, s, k=1}^{n}\left(C_{p r}^{k} C_{q s}^{k}\right)^{2} \lambda_{p}^{2} \lambda_{q}^{2}\left(\lambda_{s}^{2}+2 \lambda_{k}^{2}\right)} \tag{25}
\end{align*}
$$

Let $C^{\prime}$ be the expression in (25). Then the inequality (22) holds with $C$ replaced by $C^{\prime}$. In the case of $G=S O(3)$ equipped with the inner product $\langle X, Y\rangle=(1 / 2) \operatorname{Trace}\left(X Y^{\prime}\right)$ on $\mathfrak{g}=o(3)$, the structure constants under any orthonormal basis satisfy $C_{i j}^{k}= \pm 1$ if $i, j, k$ are distinct and $C_{i j}^{k}=0$ otherwise. Then

$$
\begin{equation*}
C^{\prime}=6 \sqrt{\lambda_{1}^{4}\left(\lambda_{2}^{2}+\lambda_{3}^{2}\right)+\lambda_{2}^{4}\left(\lambda_{1}^{2}+\lambda_{3}^{2}\right)+\lambda_{3}^{4}\left(\lambda_{1}^{2}+\lambda_{2}^{2}\right)+3 \lambda_{1}^{2} \lambda_{2}^{2} \lambda_{3}^{2}} . \tag{26}
\end{equation*}
$$

## References

[1] V.I. Arnold. Mathematical methods of classical mechanics. (1978) Springer-Verlag.
[2] S. Helgason. Differential geometry, Lie groups, and symmetric spaces. (1978) Academic Press.
[3] N. Ikeda and S. Watanabe. Stochastic differential equations and diffusion processes, 2nd ed. (1989) North-Holland-Kodansha.
[4] M. Liao. Random motion of a rigid body. J. of Theor. Prob. 10 (1997), 201-211.


[^0]:    ${ }^{1}$ THE PAPER WAS COMPLETED AT AND WAS SUPPORTED BY NANKAI UNIVERSITY

