Stein operators for variables form the third and fourth Wiener chaoses

Robert E. Gaunt

Abstract

Let \( Z \) be a standard normal random variable and let \( H_n \) denote the \( n \)-th Hermite polynomial. In this note, we obtain Stein equations for the random variables \( H_3(Z) \) and \( H_4(Z) \), which represents a first step towards developing Stein’s method for distributional limits from the third and fourth Wiener chaoses. Perhaps surprisingly, these Stein equations are fifth and third order linear ordinary differential equations, respectively. As a warm up, we obtain a Stein equation for the random variable \( aZ^2 + bZ + c, \ a, b, c \in \mathbb{R} \), which leads us to a Stein equation for the non-central chi-square distribution. We also provide a discussion as to why obtaining Stein equations for \( H_n(Z), \ n \geq 5 \), is more challenging.

Keywords: Stein’s method; normal distribution; Hermite polynomial; Wiener chaos; non-central chi-square distribution

AMS 2010 Subject Classification: Primary 60F05; 60G15; Secondary 60H07

1 Introduction

1.1 Background

In 1972, Stein [38] introduced a powerful technique for deriving distributional bounds for normal approximation. Stein’s method for normal approximation rests on the following characterisation of the normal distribution: \( W \sim N(0,1) \) if and only if

\[
\mathbb{E}[f'(W) - Wf(W)] = 0 \quad (1.1)
\]

for all real-valued absolutely continuous functions \( f \) such that \( \mathbb{E}|f'(Z)| < \infty \) for \( Z \sim N(0,1) \). This characterisation leads to the so-called Stein equation:

\[
f'(x) - xf(x) = h(x) - Nh, \quad (1.2)
\]

where \( Nh \) denotes \( \mathbb{E}h(Z) \) for \( Z \sim N(0,1) \), and the test function \( h \) is real-valued. It is straightforward to verify that \( f(x) = e^{x^2/2} \int_{-\infty}^{x} |h(t) - Nh|e^{-t^2/2} \, dt \) solves (1.2), and bounds on the solution and its derivatives in terms of the test function \( h \) and its derivatives are given in [9, 11]. Evaluating both sides of (1.2) at a random variable \( W \) and taking expectations gives

\[
\mathbb{E}[f'(W) - Wf(W)] = \mathbb{E}h(W) - Nh. \quad (1.3)
\]

Thus, the problem of bounding the quantity \( |\mathbb{E}h(W) - Nh| \) has been reduced to bounding the left-hand side of (1.3). For a detailed account of the method see the book [9].

In recent years, one of the most significant applications of Stein’s method for normal approximation has been to Gaussian analysis on Wiener space. This body of research was initiated by [28], in which Stein’s method and Malliavin calculus are combined to derive a quantitative fourth moment theorem for the normal approximation of a sequence of random variables living in a fixed Wiener chaos. A detailed account of normal approximation by the Malliavin-Stein method can be found in the book [30].
One of the advantages of Stein’s method is that the above procedure can be extended to treat many other distributional approximations; examples include the Poisson [8], gamma [21, 27], exponential [7, 35] and variance-gamma distributions [16]. The Malliavin-Stein method is also applicable to other limits, such as the multivariate normal [31], exponential and Pearson families [13, 14], centered gamma [12, 28], variance-gamma [15], linear combinations of centered chi-square random variables [2, 1, 5, 32], as well as a large family of distributions, such as the uniform, log-normal and Pareto distributions, that satisfy a diffusive assumption [25]. These works have resulted in analogues of the celebrated fourth moment theorem for non-normal limits.

However, despite these advances, there are many important distributional limits that fall outside the current state of the art of the Malliavin-Stein method. As noted by [34], an important class of limits for which little is understood are those of the type \( P \) due to their fundamental role in Gaussian analysis and Malliavin calculus.

Let \( H_n(x) \) be the \( n \)-th Hermite polynomial, defined by \( H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} (e^{-x^2/2}) \), \( n \geq 1 \), and \( H_0(x) = 1 \). The first six Hermite polynomials are then
\[
H_1(x) = x, \quad H_2(x) = x^2 - 1, \quad H_3(x) = x^3 - 3x, \quad H_4(x) = x^4 - 6x^2 + 3, \\
H_5(x) = x^5 - 10x^3 + 15x, \quad H_6(x) = x^6 - 15x^4 + 45x^2 - 15.
\]

In this note, we consider the problem of extending Stein’s method to the random variables \( H_n(Z) \). Of course, the case \( n = 1 \) is very well understood. The case \( n = 2 \) corresponds to the centered chi-square random variable \( Z^2 - 1 \). This is a special case of the centered gamma distribution, for which the Malliavin-Stein method is also highly tractable; see [3, 4, 12, 28]. As with Stein’s method for normal approximation, at the heart of Stein’s method for \( H_2(Z) = Z^2 - 1 \) is the Stein equation
\[
2(1 + x)f'(x) - xf(x) = h(x) - N_2 h,
\]
where \( N_i h \) denotes the quantity \( \mathbb{E}[h(H_i(Z))] \), \( i \geq 1 \). The Stein equation (1.4) follows by applying a simple translation to the classic gamma Stein equation of [10, 27]. Estimates for the solution of the gamma Stein equation [11, 12, 21, 27] can then be used to bound the solution of (1.4) and its derivatives. The problem of approximating a random variable \( W \) by \( H_2(Z) \) is thus reduced to the tractable one of bounding the quantity \( \mathbb{E}[2(1 + W)f'(W) - Wf(W)] \).

Recently, [1] obtained a Stein equation for random variables of the form \( F_\infty = \sum_{i=1}^q \alpha_i (Z_i^2 - 1) \), where the \( \alpha_i \) are real-valued constants and \( Z_1, \ldots, Z_q \) are independent \( N(0,1) \) random variables. Their Stein equation for \( F_\infty \) was \( q \)-th order if the \( \alpha_i \) are all distinct, and to date no bound exists for the solution of the Stein equation when \( q \geq 3 \); the case \( q = 2 \) corresponds to the variance-gamma Stein equation, and bounds from [11, 16] can be used. Despite not being able to bound the solution of the Stein equation in general, their Stein equation motivated a Stein kernel for \( F_\infty \) that was used in [2] to bound the 2-Wasserstein distance between a random variable \( F \) belonging to the second Wiener chaos and the limit \( F_\infty \).

### 1.2 Summary of results

In this note, we consider the problem of extending Stein’s method to \( H_n(Z) \), \( n \geq 3 \). In particular, we study the problem of finding Stein equations for \( H_n(Z) \), \( n \geq 3 \). Indeed, our main results are the following Stein equations (see Propositions 2.4 and 2.5) for \( H_3(Z) \) and \( H_4(Z) \):
\[
486(4 - x^2)f^{(5)}(x) - 486xf^{(4)}(x) - 27(8 - x^2)f^{(3)}(x) \\
+ 99xf''(x) + 6f'(x) - xf(x) = h(x) - N_3 h,
\]
(1.5)
and

\[ 192(x+6)(3-x)f^{(3)}(x) + 16(x+3)(x-12)f''(x) + 4(11x+6)f'(x) - xf(x) = h(x) - N_4 h. \]

(1.6)

The first striking feature of the Stein equations (1.5) and (1.6) is that they are fifth and third order linear differential equations, respectively. There is a dramatic increase in complexity from the Stein equations (1.2) and (1.4) for \( H_1(Z) \) and \( H_2(Z) \) to the Stein equation for \( H_3(Z) \). Such higher order Stein operators were uncommon in the literature, with the exception of [22], but following the introduction of second order operators by [16, 36, 37] and a \( n \)-th order operator for the product of \( n \) independent \( N(0,1) \) random variables [17], have appeared in several recent papers, and an overview is given in [20]. It is also perhaps surprising that the Stein equation for \( H_4(Z) \) is of lower order than the Stein equation for \( H_3(Z) \). We shed some light on this feature in Remark 2.9.

As (1.5) and (1.6) are differential equations of order greater than two, neither can be readily obtained by the standard generator [6, 23] or density approaches [26, 39]. To obtain the Stein equations (1.5) and (1.6) we employ a recent technique that was introduced in Section 4.3 of [20]. The same technique could in principle be used to obtain Stein equations for \( H_n(Z) \), \( n \geq 5 \), although this would most likely lead to more involved calculations; see Section 2.3 for further comments. In Proposition 2.1 we also use the technique to obtain a Stein equation for the random variable \( aZ^2 + bZ + c \). That Stein equation is new and serves as a simple illustration of the technique. Moreover, this result leads to a Stein equation for the non-central chi square distribution (see Proposition 2.3 and equation (2.15)). This Stein equation is a second order differential equation and it would be relatively straightforward to solve the equation and bound its solution, which would yield the first two ingredients of Stein’s method for non-central chi-square approximation.

Solving and bounding the solutions of (1.5) and (1.6) is, however, far more challenging and we were unable to make progress here. Despite recent advances on this aspect of Stein’s method [11], such results are beyond the current state of the art. However, the Stein equations (1.5) and (1.6) may yield insight into convergence in the third and fourth Wiener chaoses, in the same way that the Stein equation for linear combinations of chi-square random variables of [11] provided important motivation for approximation of random variables in the second Wiener chaos [2]. As this note is devoted to the problem of finding Stein equations for \( H_n(Z) \), we shall not explore the techniques that may be required for such an investigation, such as a suitable analogue of the Stein kernel (as used by [21]) or iterated gamma operators [29] that were used by [15] in developing the Malliavin-Stein method for variance-gamma approximation. Nevertheless, this work can be viewed as a first step towards the fundamental problem of understanding Stein operators for general functionals of normal random variables, and to the development of Stein’s method for limit distributions from the third and fourth Wiener chaoses.

2 Stein equations for \( P(Z) \)

Throughout this section, we shall let \( C^k(I) \) denote the space of functions that are \( k \)-times differentiable on the interval \( I \subseteq \mathbb{R} \). It is understood that, for a function \( h \), \( h^{(0)} \equiv h \). The supremum norm of \( f : I \rightarrow \mathbb{R} \) is denoted by \( \|f\| = \|f\|_\infty = \sup_{x \in I} |f(x)| \).

2.1 Stein equations for \( aZ^2 + bZ + c \) and the non-central chi-square distribution

The following proposition gives a characterising equation for the random variable \( aZ^2 + bZ + c \), \( a, b, c \in \mathbb{R} \), which leads to a corresponding Stein equation. The proof provides a simple illustration of a recent technique of [20].
Proposition 2.1. Let $X = aZ^2 + bZ + c$, where $a, b, c \in \mathbb{R}$. Let $f \in C^2((\alpha, \beta))$, where $(\alpha, \beta) \subset \mathbb{R}$ is the range of the quadratic $ax^2 + bx + c$. Suppose further that $\mathbb{E}|X^k| < \infty$ and $\mathbb{E}|f^{(k)}(X)| < \infty$ for all $k = 0, 1, 2$. Then

$$\mathbb{E}\left[(ab^2 + 4a^2(X - c))f''(X) + (2a^2 - b^2 - 4a(X - c))f'(X) + (X - c - a)f(X)\right] = 0. \quad (2.7)$$

Proof. To simplify the calculations, we find a characterising equation for the random variable $X = aZ^2 + bZ + c$ and then apply a translation to deduce a characterising equation for $Z^2 + bZ + c$. Recall from (1.1) that, for suitable $g$, $\mathbb{E}Zg(Z) = \mathbb{E}g'(Z)$. Then, since $\frac{d}{dx}(x^{k-1}f(w)) = (2ax^k + bx^{k-1})f'(w) + (k - 1)x^{k-2}f(w)$ for $k \geq 1$, where $w = ax^2 + bx$, we have that

$$\mathbb{E}Z^k f(W) = 2a\mathbb{E}Z^k f'(W) + b\mathbb{E}Z^{k-1} f'(W) + (k - 1)\mathbb{E}Z^{k-2} f(W), \quad k \geq 1. \quad (2.8)$$

Using (2.8) gives that

$$\mathbb{E}Wf(W) = a\mathbb{E}Z^2 f(W) + b\mathbb{E}f(W) = 2a\mathbb{E}Z f'(W) + ab\mathbb{E}Z f'(W) + a\mathbb{E}f(W) + 2ab\mathbb{E}Z f'(W) + b^2\mathbb{E}f'(W) = 2a\mathbb{E}Wf'(W) + b^2\mathbb{E}f'(W) + a\mathbb{E}f(W) + ab\mathbb{E}f'(W). \quad (2.9)$$

Another application of (2.8) gives that

$$\mathbb{E}Wf(W) = 2a\mathbb{E}W f'(W) + b^2\mathbb{E}f'(W) + a\mathbb{E}f(W) + 2a^2b\mathbb{E}Z f''(W) + ab^2\mathbb{E}f''(W). \quad (2.10)$$

On rearranging (2.10) we obtain

$$2a^2b\mathbb{E}Z f''(W) = \mathbb{E}W f(W) - 2a\mathbb{E}W f'(W) - b^2\mathbb{E}f'(W) - a\mathbb{E}f(W) - ab^2\mathbb{E}f''(W), \quad (2.11)$$

and from (2.9) applied to $f'$ we obtain

$$2a^2b\mathbb{E}Z f''(W) = 2a\mathbb{E}W f'(W) - 4a^2\mathbb{E}W f''(W) - 2ab^2\mathbb{E}f''(W) - 2a^2\mathbb{E}f'(W). \quad (2.12)$$

Combining (2.11) and (2.12) gives

$$\mathbb{E}\left[(ab^2 + 4a^2W)f''(W) + (2a^2 - b^2 - 4aW)f'(W) + (W - a)f(W)\right] = 0. \quad (2.7)$$

Finally, applying a simple translation yields (2.7), as required. \qed

Remark 2.2. The first part of the proof involved obtaining equation (2.7). All terms on the right hand side of (2.7) are expressed in terms of $W$, except for the final term $ab\mathbb{E}Z f'(W)$. In the next part of the proof, we used the technique of [20] to obtain another expression, equation (2.11), for $\mathbb{E}Z f''(W)$ in terms of expectations involving $W$. The ‘awkward’ $\mathbb{E}Z f''(W)$ term can then be cancelled out. We use this basic strategy, together with more involved calculations, to obtain our Stein equations for $H_3(Z)$ and $H_4(Z)$.

The probability density function of the non-central chi-square distribution with $k > 0$ degrees of freedom and non-centrality parameter $\lambda > 0$ is given by

$$p(x) = \frac{1}{2}e^{-(x+\lambda)/2} \left(\frac{x}{\lambda}\right)^{k/4-1/2} I_{k-1}(\sqrt{\lambda}x), \quad x > 0, \quad (2.13)$$

where $I_\nu(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!(\nu+k+1)} I_{\nu+k+1}(\frac{x^2}{2})$ is a modified Bessel function of the first kind (see [33]). For distributional properties see [21]. If the random variable $Y$ has density (2.13), we write $Y \sim \chi^2_k(\lambda)$. The non-central chi-square distribution $\chi^2_k(\lambda)$ is equal in distribution to $\sum_{i=1}^{k} X_i^2$, where $X_1, \ldots, X_k$ are independent $N(\mu_i, 1)$ random variables, $1 \leq i \leq k$, and $\lambda = \sum_{i=1}^{k} \mu_i^2$. With the aid of Proposition 2.1 we can obtain the following proposition.
Proposition 2.3. Let $X$ have the $\chi^2_k(\lambda)$ distribution. Let $f \in C^2((0, \infty))$ be such that $\mathbb{E}|Xf^{(j)}(X)| < \infty$, $j = 0, 1, 2$, and $\mathbb{E}|f^{(k)}(X)| < \infty$, $k = 0, 1$. Then
\begin{equation}
\mathbb{E}[4Xf''(X) + (2k - 4X)f'(X) + (X - k - \lambda)f(X)] = 0.
\end{equation}

Proposition 2.3 leads to the following Stein equation for the $\chi^2_k(\lambda)$ distribution:
\begin{equation}
4xf''(x) + (2k - 4x)f'(x) + (x - k - \lambda)f(x) = h(x) - \chi^2_k(\lambda),
\end{equation}
where $\chi^2_k(\lambda)$ denotes $\mathbb{E}h(Y)$ for $Y \sim \chi^2_k(\lambda)$.

Proof. Let $X_i \sim \chi^2_1(\mu_i), 1 \leq i \leq k$, be independent, and suppose that $\lambda = \sum_{i=1}^k \mu_i^2$. Then $X_i \overset{d}{=} (Z + \mu_i)^2 = Z^2 + 2\mu_i Z + \mu_i^2$, where $Z \sim N(0, 1)$. Setting $a = 1, b = 2\mu_i$ and $c = \mu_i^2$ in (2.7) gives that $\mathbb{E}[A_{X_i}f(X_i)] = 0, 1 \leq i \leq k$, where the Stein operator $A_{X_i}$ is given by
\begin{equation}
A_{X_i}f(x) = 4xf''(x) + (2 - 4x)f'(x) + (x - 1 - \mu_i^2)f(x), \quad 1 \leq i \leq k.
\end{equation}
Since the random variables $X_1, \ldots, X_k$ are independent, $Y = \sum_{i=1}^k X_i$, and the Stein operator has linear coefficients, we can deduce (2.14) form the result of Section 4.1 of [20].

2.2 Stein equations for $H_3(Z)$ and $H_4(Z)$

The following propositions lead to the Stein equations (1.5) and (1.6). Some remarks are given after the proofs.

Proposition 2.4. Let $X = H_3(Z) = Z^3 - 3Z$. Let $f \in C^5(\mathbb{R})$ and suppose $\mathbb{E}|X^2f^{(i)}(X)| < \infty$ for $i = 3, 5$, $\mathbb{E}|Xf^{(j)}(X)| < \infty$ for $j = 0, 2, 4$, and $\mathbb{E}|f^{(k)}(X)| < \infty$ for $k = 1, 3, 5$. Denote this class of functions $F_3$. Then
\begin{equation}
\mathbb{E}[486(4 - X^2)f^{(3)}(X) - 486Xf^{(4)}(X) - 27(8 - X^2)f^{(3)}(X) + 99Xf''(X) + 6f'(X) - Xf(X)] = 0.
\end{equation}

Proposition 2.5. Let $X = H_4(Z) = Z^4 - 6Z + 3$. Let $f \in C^3((-6, \infty))$ and suppose that $\mathbb{E}|X^2f^{(i)}(X)| < \infty$ for $i = 2, 3$, $\mathbb{E}|Xf^{(j)}(X)| < \infty$ for $j = 0, 1, 2, 3$, and $\mathbb{E}|f^{(k)}(X)| < \infty$ for $k = 1, 2, 3$. Denote this class of functions $F_4$. Then
\begin{equation}
\mathbb{E}[192(X + 6)(3 - X)f^{(3)}(X) + 16(X + 3)(X - 12)f''(X) + 4(11X + 6)f'(X) - Xf(X)] = 0.
\end{equation}

Proof of Proposition 2.4. Let $W = aZ^3 + bZ$. We shall later set $a = 1$ and $b = -3$. Since $\frac{d}{dx}(x^{k-1}f(w)) = (3ax^{k+1} + bx^{k-1})f'(w) + (k - 1)x^{k-2}f(w)$ for $k \geq 1$, where $w = ax^3 + bx$, we have from (1.1) that
\begin{equation}
EZ^k f(W) = 3aEZ^{k+1} f'(W) + bEZ^{k-1} f'(W) + (k - 1)EZ^{k-2} f(W), \quad k \geq 1.
\end{equation}

Using (2.18) gives that
\begin{align*}
\mathbb{E}f(W) &= aEZ^3 f(W) + bEZ f(W) \\
&= 3a^2EZ^4 f'(W) + abEZ^2 f'(W) + 2aEZ f(W) + 3abEZ^2 f'(W) + b^2EZ f'(W) \\
&= 3a^2EZ^4 f(W) + 4abEZ^2 f'(W) + 2aEZ f(W) + b^2EZ f'(W)
\end{align*}

Another application of (2.18) gives that
\begin{align*}
\mathbb{E}f(W) &= 9a^3EZ^5 f''(W) + 3a^2bEZ^3 f''(W) + 9a^2EZ^2 f'(W) + 12a^2bEZ^3 f''(W) + 4ab^2EZ f''(W) + 6a^2EZ^2 f'(W) + 2abEZ f'(W) + b^2EZ f'(W) \\
&= 9a^3EZ^5 f'(W) + 15a^2bEZ^3 f''(W) + 4ab^2EZ f''(W) + 15a^2EZ^2 f'(W) + (6ab + b^2)EZ f'(W).
\end{align*}
Following the same procedure again yields
\[
\mathbb{E}W f(W) = 27a^4\mathbb{E}Z^6 f^{(3)}(W) + 9a^3 b\mathbb{E}Z^4 f^{(3)}(W) + 36a^3 \mathbb{E}Z^3 f''(W) \\
+ 45a^3 \mathbb{E}Z^4 f^{(3)}(W) + 15a^2 b\mathbb{E}Z^2 f^{(3)}(W) + 30a^2 b\mathbb{E}Z f''(W) \\
+ 12a^2 b\mathbb{E}Z^2 f^{(3)}(W) + 4a b^2 \mathbb{E}f^{(3)}(W) \\
+ 45a^3 \mathbb{E}Z^3 f''(W) + 15a^2 b\mathbb{E}Z f''(W) + 15a^2 f'(W) + (6ab + b^2)\mathbb{E}f'(W) \\
= 27a^4\mathbb{E}Z^6 f^{(3)}(W) + 54a^3 b\mathbb{E}Z^4 f^{(3)}(W) + 27a^2 b^2\mathbb{E}Z^2 f^{(3)}(W) + 81a^3 \mathbb{E}Z^3 f''(W) \\
+ 45a^2 b\mathbb{E}Z f''(W) + 4ab^2 \mathbb{E}f^{(3)}(W) + (15a^2 + 6ab + b^2)\mathbb{E}f'(W) \\
= 27a^2\mathbb{E}W^2 f^{(3)}(W) + 4ab^2 \mathbb{E}f^{(3)}(W) + (15a^2 + 6ab + b^2)\mathbb{E}f'(W) \\
+ 81a^3 \mathbb{E}Z^3 f''(W) + 45a^2 b\mathbb{E}Z f''(W) \\
= 27a^2\mathbb{E}W^2 f^{(3)}(W) + 4ab^2 \mathbb{E}f^{(3)}(W) + (15a^2 + 6ab + b^2)\mathbb{E}f'(W) \\
+ 81a^2 \mathbb{E}W f''(W) - 36a^2 b\mathbb{E}Z f''(W).
\]  
(2.19)

We now note that
\[
\mathbb{E}Z f''(W) = 3a\mathbb{E}Z^2 f^{(3)}(W) + b\mathbb{E}f^{(3)}(W) \\
= 9a^2 \mathbb{E}Z^3 f^{(4)}(W) + 3ab\mathbb{E}Z f^{(4)}(W) + (3a + b)\mathbb{E}f^{(3)}(W) \\
= 9a\mathbb{E}W f^{(4)}(W) - 6ab\mathbb{E}Z f^{(4)}(W) + (3a + b)\mathbb{E}f^{(3)}(W). 
\]  
(2.20)

On substituting (2.20) into (2.19) we obtain
\[
\mathbb{E}W f(W) = 27a^2\mathbb{E}W^2 f^{(3)}(W) + 4ab^2 \mathbb{E}f^{(3)}(W) + (15a^2 + 6ab + b^2)\mathbb{E}f'(W) \\
+ 81a^2 \mathbb{E}W f''(W) - 324a^3 b\mathbb{E}W f^{(4)}(W) - (108a^3 b + 36a^2 b^2)\mathbb{E}f^{(3)}(W) \\
+ 216a^3 b^2 \mathbb{E}Z f^{(4)}(W). 
\]  
(2.21)

On rearranging (2.21) we obtain
\[
216a^3 b^2 \mathbb{E}Z f^{(4)}(W) = \mathbb{E}W f(W) + 324a^3 b\mathbb{E}W f^{(4)}(W) \\
+ (108a^3 b + 36a^2 b^2 - 4ab^3)\mathbb{E}f^{(3)}(W) - 27a^2\mathbb{E}W^2 f^{(3)}(W) \\
- 81a^2 \mathbb{E}W f''(W) - (15a^2 + 6ab + b^2)\mathbb{E}f'(W), 
\]  
(2.22)

and from (2.19) applied to $f''$ we obtain
\[
216a^3 b^2 \mathbb{E}Z f^{(4)}(W) = 6ab\left\{27a^2\mathbb{E}W^2 f^{(5)}(W) + 4ab^3 \mathbb{E}f^{(5)}(W) + 81a^2 \mathbb{E}W f^{(4)}(W) \\
+ (15a^2 + 6ab + b^2)\mathbb{E}f^{(3)}(W) - \mathbb{E}W f''(W)\right\} \\
= 162a^3 b\mathbb{E}W^2 f^{(3)}(W) + 24a^2 b^4 \mathbb{E}f^{(5)}(W) + 486a^3 b\mathbb{E}W f^{(4)}(W) \\
+ (90a^3 b + 36a^2 b^2 + 6ab^3)\mathbb{E}f^{(3)}(W) - 6ab\mathbb{E}W f''(W). 
\]  
(2.23)

Combining (2.22) and (2.23) now yields
\[
\mathbb{E}\left[(162a^3 b W^2 + 24a^2 b^4) f^{(5)}(W) + 162a^3 b W f^{(4)}(W) + (27a^2 W^2 - 18a^3 b + 2ab^3) f^{(3)}(W) \\
+ (81a^2 - 6ab) W f''(W) + (15a^2 + 6ab + b^2) f'(W) - W f(W)\right] = 0. 
\]

Finally, on setting $a = 1$ and $b = -3$ we obtain (2.16), as required.

Proof of Proposition 2.4 To simplify the calculations, we find a characterising equation for the random variable $W = aZ^4 + bZ^2$ and then apply a translation to deduce a characterising equation for the
random variable $aZ^4 + bZ^2 + c$. We shall then set $a = 1$, $b = -6$ and $c = 3$. Since $\frac{d}{dw}(x^{k-1}f(w)) = (4ax^{k+2} + 2bx^k)f'(w) + (k-1)x^{k-2}f(w)$ for $k \geq 1$, where $w = ax^4 + bx^2$, we have from (11.1) that
\[
\mathbb{E}Z^k f(W) = 4a\mathbb{E}Z^{k+2} f'(W) + 2b\mathbb{E}Z^k f'(W) + (k-1)\mathbb{E}Z^{k-2} f(W), \quad k \geq 1. \tag{2.24}
\]
Using (2.24) gives that
\[
\mathbb{E}W f(W) = a\mathbb{E}Z^4 f(W) + b\mathbb{E}Z^2 f(W)
= 4a^2\mathbb{E}Z^6 f'(W) + 2ab\mathbb{E}Z^4 f'(W) + 3a\mathbb{E}Z^2 f(W)
+ 4ab\mathbb{E}Z^4 + 2b^2\mathbb{E}Z^2 f'(W) + b\mathbf{f}(W)
= 4a^2\mathbb{E}Z^6 f'(W) + 6ab\mathbb{E}Z^4 f'(W) + 3a\mathbb{E}Z^2 f(W) + 2b^2\mathbb{E}Z^2 f'(W) + b\mathbf{f}(W).
\]
Another application of (2.24) gives that
\[
\mathbb{E}W f(W) = 16a^3\mathbb{E}Z^8 f''(W) + 8a^2b\mathbb{E}Z^6 f''(W) + 20a^2\mathbb{E}Z^4 f'(W) + 24a^2b\mathbb{E}Z^6 f''(W)
+ 12ab^2\mathbb{E}Z^4 f''(W) + 18ab\mathbb{E}Z^2 f'(W) + 8ab^2\mathbb{E}Z^4 f''(W) + 4b^3\mathbb{E}Z^2 f''(W)
+ 2b^2\mathbf{f}'(W) + 12a^2\mathbb{E}Z^4 f'(W) + 6ab\mathbb{E}Z^2 f'(W) + 3a\mathbf{f}(W) + b\mathbf{f}(W)
= 16a^3\mathbb{E}Z^8 f''(W) + 32a^2b\mathbb{E}Z^6 f''(W) + 20ab^2\mathbb{E}Z^4 f''(W) + 4b^3\mathbb{E}Z^2 f''(W)
+ 32a^2\mathbb{E}Z^4 f'(W) + 24ab\mathbb{E}Z^2 f'(W) + 2b^2\mathbf{f}'(W) + (3a + b)\mathbf{f}(W)
= 16a\mathbb{E}W^2 f''(W) + 4b^2\mathbb{E}W f''(W) + 32a\mathbb{E}W f'(W) + 2b^2\mathbf{f}'(W)
+ (3a + b)\mathbf{f}(W) - 8ab\mathbb{E}Z^2 f'(W). \tag{2.25}
\]
We now note that
\[
\mathbb{E}Z^2 f'(W) = 4a\mathbb{E}Z^4 f''(W) + 2b\mathbb{E}Z^2 f''(W) + \mathbf{f}'(W)
= 4\mathbb{E}W f''(W) - 2b\mathbb{E}Z^2 f''(W) + \mathbf{f}'(W). \tag{2.26}
\]
On substituting (2.26) into (2.25) we obtain
\[
\mathbb{E}W f(W) = 16a\mathbb{E}W^2 f''(W) + (4b^2 - 32ab)\mathbb{E}W f''(W) + 32a\mathbb{E}W f'(W)
+ (2b^2 - 8ab)\mathbf{f}'(W) + (3a + b)\mathbf{f}(W) + 16ab^2 Z^2 f''(W). \tag{2.27}
\]
On rearranging (2.27) we obtain
\[
16ab^2\mathbb{E}Z^2 f''(W) = \mathbb{E}W f(W) - 16a\mathbb{E}W^2 f''(W) + (32ab - 4b^2)\mathbb{E}W f''(W)
- 32a\mathbb{E}W f'(W) + (8ab - 2b^2)\mathbf{f}'(W) - (3a + b)\mathbf{f}(W), \tag{2.28}
\]
and from (2.25) applied to $f'$ we obtain
\[
16ab^2\mathbb{E}Z^2 f''(W) = 2b\left\{-\mathbb{E}W f'(W) + 16a\mathbb{E}W^2 f^{(3)}(W) + 4b^2\mathbb{E}W f^{(3)}(W)
+ 32a\mathbb{E}W f''(W) + 2b^2\mathbf{f}'(W) + (3a + b)\mathbf{f}'(W)\right\}
= -2b\mathbb{E}W f'(W) + 32a\mathbb{E}W^2 f^{(3)}(W) + 8b^3\mathbb{E}W f^{(3)}(W)
+ 64ab\mathbb{E}W f''(W) + 4b^3\mathbf{f}'(W) + (6ab + 2b^2)\mathbf{f}'(W). \tag{2.29}
\]
Combining (2.28) and (2.29) now yields
\[
\mathbb{E}\left[(32abW^2 + 8b^3W) f^{(3)}(W) + (16aW^2 + (4b^2 + 32ab)W + 4b^3) f''(W)
+ ((32a - 2b)W + 4b^2 - 2ab)f'(W) + (3a + b - W) f(W)\right] = 0.
\]
By applying a translation, we deduce the following characterising equation for the random variable $X = aZ^4 + bZ^2 + c$:

$$
\mathbb{E}[(32ab(X-c)^2 + 8b^3(X-c))f^{(3)}(X) + (16a(X-c)^2 + (4b^2 + 32ab)(X-c) + 4b^3)f''(X)
+ ((32 - 2b)(X-c) + 4b^2 - 2ab)f'(X) - (X - 3a - b - c)f(X)] = 0.
$$

Finally, on setting $a = 1$, $b = -6$ and $c = 3$ we obtain (2.17), as required. \qed

**Remark 2.6.** In the proofs of Propositions 2.4 and 2.5 we actually obtained characterising equations for the more general random variables $aZ^3 + bZ$ and $aZ^4 + bZ^2 + c$. The reason for working with these more general random variables is that it enables one to combine terms of the type $\mathbb{E}Z^k f^{(j)}(W)$ to form terms of the type $\mathbb{E}W^k f^{(j)}(W)$. For example, a priori it is not immediately clear whether the term $\mathbb{E}Z^3 f^{(3)}(W)$ should be combined to form $\mathbb{E}W^3 f^{(3)}(W)$ or $\mathbb{E}W f^{(3)}(W)$, but when multiplied by an appropriate constant in terms of $a$ and $b$ this become clear. Indeed, a reasonable amount of trial and error was used in deriving the characterising equations for $aZ^3 + bZ$ and $aZ^4 + bZ^2 + c$, and the introduction of the constants $a$, $b$ and $c$ greatly helped in this regard.

It is natural to ask for Stein operators for the more general polynomials $P_3(Z) = aZ^3 + bZ^2 + cZ + d$ and $P_4(Z) = aZ^4 + bZ^3 + cZ^2 + dZ + e$. However, this leads to non-trivial further complications. A discussion on why there is an increase in difficulty for the general quartic $P_4(Z)$ is given in Section 2.3. That there is an increase in difficulty for the general quartic $P_4(Z)$ is easy to see: a Stein equation for $P_4(Z)$ must include the fifth order Stein equation for $aZ^3 + bZ$ as a special case.

**Remark 2.7.** It should be possible, via the density method [38, 20] for example, to obtain Stein equations for $H_3(Z)$ and $H_4(Z)$ that are differential equations with order less than five and three respectively, although this would most likely come at the cost of more complicated coefficients than the polynomial coefficients of (1.5) and (1.6). We strongly believe that (1.5) and (1.6) are the minimal order Stein equations with polynomial coefficients for $H_3(Z)$ and $H_4(Z)$, in the sense that there do not exist Stein equations for $H_3(Z)$ and $H_4(Z)$ that are polynomial coefficient differential equations of order less than five and three, respectively. We are unable to prove this, but a future work will investigate this assertion for polynomials of a fixed low degree.

**Remark 2.8.** Let us now provide some insight regarding the orders of the Stein operators obtained in this note. It is perhaps surprising that our Stein operator for the non-central chi-square distribution is of second order, whereas the classical Stein operator for the centered chi-square distribution is first order. This is a consequence of the duality between Stein operators and differential equations satisfied by the p.d.f. of the distribution of interest. Roughly speaking, if the p.d.f. satisfies a $k$-th order differential equation with polynomial coefficients, then, under fairly mild assumptions, there will exist a $k$-th order Stein operator for the distribution (see [14, 20, 26] for more details). The presence of the modified Bessel function of the first kind (which is known to satisfy a second order differential equation with polynomial coefficients) means that the p.d.f. of the non-central chi-square distribution satisfies a second order differential equation, and so by duality we arrive at a second order Stein operator. In contrast, the p.d.f. of the centered chi-square distribution is given in terms of elementary functions and satisfies a first order differential equation.

We now provide two pieces of intuition as to why our Stein operator for $H_3(Z)$ is of greater order than our Stein operator for $H_4(Z)$. One is again due to the duality between p.d.f.s and Stein operators. The p.d.f.s for $H_3(Z)$ and $H_4(Z)$ can be obtained via elementary, but tedious, calculations. We omit the details, but make the observation that the p.d.f.s are defined piecewise, which helps explain why the Stein operators are of order greater than one (see pp. 574–575 of [37] for details regarding the derivation of Stein equations for the piecewisely defined Laplace distribution). To see this, note that in order to find the p.d.f.s of $H_3(Z)$ and $H_4(Z)$ via the usual method for finding p.d.f.s of functions of random variables one must solve the equations $z^3 - 3z = x$ and $z^4 - 6z^2 + 3 = x$, respectively, for $z$, and the nature of the
solutions will change as \( x \) varies over the real line. The solution of \( z^3 - 3z = x \) is more complicated than that of \( z^4 - 6z^2 + 3 = x \), as it requires the use of the cubic formula rather than the quadratic formula, and consequently the p.d.f. of \( H_3(Z) \) will take a more complicated form than that of \( H_4(Z) \). It therefore seems plausible that the p.d.f. of \( H_3(Z) \) will only satisfy, in a piecewise sense, a polynomial coefficient differential equation that is of higher order than that of \( H_4(Z) \), meaning by duality that the Stein operator will also be of higher order.

Another piece of intuition is that obtaining a Stein equation for \( H_4(Z) \) is no more difficult than obtaining one for \( Z^4 - 6Z^2 \). Now \( Z^2 \) has the \( \chi^2_{(1)} \) distribution, for which the classical Stein equation is first order. Therefore finding a Stein equation for \( Z \) seems plausible that the p.d.f. as defined in Proposition 2.4, then \( W \) cannot use the standard method for proving sufficiency of Stein characterisations (see, for example, the proof of Lemma 2.1 of [9]). Moreover, the law of \( H_n(Z) \), \( n \geq 3 \), is not determined by its moments. Therefore, we cannot appeal to Proposition 4.8 of [20] to prove sufficiency. (Note that one can perfectly well apply Stein’s method without establishing sufficiency; see, for example, [36].) However, we can use the argument used to prove Proposition 4.8 of [20] to obtain a weak form of sufficiency. If for a random variable \( W_3 \), which shares its first five moments with \( H_3(Z) \), the expectation (2.10) holds for all \( f \in \mathcal{F}_3 \), as defined in Proposition 2.4, then \( W_3 \) must share all moments with \( H_3(Z) \), and if for a random variable \( W_4 \), which shares its first three moments with \( H_4(Z) \), the expectation (2.10) holds for all \( f \in \mathcal{F}_4 \) then \( W_4 \) must share all moments with \( H_4(Z) \).

### 2.3 Stein equations for \( H_n(Z) \), \( n \geq 5 \)

The next obvious problems are to find Stein equations for \( P_3(Z) \), where \( P_3 \) is a general cubic, and \( H_n(Z) \), \( n \geq 5 \). Based on our Stein equations for \( aZ^2 + bZ + c \), \( H_3(Z) \) and \( H_4(Z) \), we expect that the order of increasing difficulty is: \( P_3(Z) \), \( H_6(Z) \) and then \( H_5(Z) \). Before tackling the general case of \( P_3(Z) \), it may be worth working out the special case \( W = aZ^3 + bZ^2 \). Even here one faces additional difficulties. Unlike in our proofs of Propositions 2.1, 2.4 and 2.3, it is not possible for any \( i \geq 1 \) to find a \( j \geq 1 \) such that \( EZ^i f(W) = EZ^j f^{(j)}(W) + R \), where \( R \) is an expression that can be solely written in terms of expectations involving \( W \). (This author has encountered this same problem whilst trying to obtain Stein operators for various product distributions, and considers it to be important for methods to be developed to overcome this difficulty.) This means that a more sophisticated implementation of the technique of [20] may be needed to deal with the ‘awkward’ expressions \( EZ^i f^{(k)}(W) \). Even with a more ingenious implementation of this technique, the complexity of the calculations would most likely mean that the approach would become intractable for \( H_7(Z) \), \( H_8(Z) \) and beyond. Obtaining Stein equations for such random variables will therefore most likely require the introduction of new techniques.

The following observation may be useful in the quest for Stein equations for \( H_n(Z) \), \( n \geq 5 \). In Table 1, we record the coefficient of the highest order derivative of the Stein equation for \( H_n(Z) \) for \( n = 1, \ldots, 4 \), and note its connection to the local maxima and minima of the Hermite polynomials. This allows us to conjecture the corresponding terms for the Stein equations for \( H_5(Z) \) and \( H_6(Z) \). For \( H_5(Z) \) we expect the leading coefficient to be

\[
\left( x - 4\sqrt{6(3 + \sqrt{6})} \right) \left( x - 4\sqrt{6(3 + \sqrt{6})} \right) \left( x + 4\sqrt{6(3 + \sqrt{6})} \right) \left( x + 4\sqrt{6(3 - \sqrt{6})} \right)
\]

\[
= x^4 - 576x^2 + 27648,
\]

and for \( H_6(Z) \) we expect it to be

\[
\left( x - 20(\sqrt{10} - 2) \right) \left( x + 20(2 + \sqrt{10}) \right) \left( x + 15 \right) = x^3 + 95x^2 - 1200x - 3600.
\]
In the quest for Stein equations for $H_n(Z)$, $n \geq 5$, it would also be useful to have a conjecture on the order of the Stein equations. This author does not have a conjecture on the exact order, but does strongly suspect that the Stein equations for $H_5(Z)$ and $H_6(Z)$ will be of at least ninth and sixth order, respectively.

Table 1: Local maxima and minima of Hermite polynomials $H_n(x)$ and coefficient of highest order derivative of the Stein equation for $H_n(Z)$.

| $H_n(x)$ | Local maxima | Local minima | Coefficient |
|----------|--------------|--------------|-------------|
| $H_1(x)$ | -            | -            | 1           |
| $H_2(x)$ | -            | -1           | $x$         |
| $H_3(x)$ | 2            | -2           | $x^2 - 4$   |
| $H_4(x)$ | 3            | -6 (twice)   | $(x + 6)(x - 3)$ |
| $H_5(x)$ | $4\sqrt{6(3 + \sqrt{6})}$, $4\sqrt{6(3 - \sqrt{6})}$ | $-4\sqrt{6(3 + \sqrt{6})}$, $-4\sqrt{6(3 - \sqrt{6})}$ | ? |
| $H_6(x)$ | $20(\sqrt{10} - 2)$ (twice) | $-20(2 + \sqrt{10})$ (twice), $-15$ | ? |

Acknowledgements

The author is supported by a Dame Kathleen Ollerenshaw Research Fellowship. The author would like to thank Ivan Nourdin for suggesting the problem of finding Stein equations for $H_n(Z)$, $n \geq 3$. The author would also like to thank the reviewer for their comments and suggestions which lead to an improved article, and is grateful to Ehsan Azmoodeh and Dario Gasbarra for bringing to my attention a mistake in the Stein equation for $H_4(Z)$ that was present in an earlier version.

References

[1] Arras, B., Azmoodeh, E., Poly, G. and Swan, Y. Stein characterizations for linear combinations of gamma random variables. arXiv:1709:01161, 2017.
[2] Arras, B., Azmoodeh, E., Poly, G. and Swan, Y. A bound on the 2-Wasserstein distance between linear combinations of independent random variables. To appear in Stoch. Proc. Appl., 2018+.
[3] Azmoodeh, E., Campese, S. and Poly, G. Fourth Moment Theorems for Markov Diffusion Generators. J. Funct. Anal. 266 (2014), pp. 2341–2359.
[4] Azmoodeh, E., Malicet, D., Mijoule, G. and Poly, G. Generalization of the Nualart-Peccati criteria. Ann. Probab. (2016), pp. 924–954.
[5] Azmoodeh, E., Peccati, G. and Poly, G. Convergence towards linear combinations of chi-squared random variables: a Malliavin-based approach. Séminaire de Probabilités XLVII (special volume in memory of Marc Yor) (2015), pp. 339–367.
[6] Barbour, A. D. Stein’s method for diffusion approximations. Probab. Theory Rel. 84 (1990), pp. 297–322.
[7] Chatterjee, S., Fulman, J. and Röllin, A. Exponential approximation by Stein’s method and spectral graph theory. ALEA Lat. Am. J. Probab. Math. Stat. 8 (2011), pp. 197–223.
[8] Chen, L. H. Y. Poisson approximation for dependent trials. Ann. Probab. 3 (1975), pp. 534–545.
[9] Chen, L. H. Y., Goldstein, L. and Shao, Q.–M. Normal Approximation by Stein’s Method. Springer, 2011.
[10] Diaconis, P. and Zabell, S. Closed Form Summation for Classical Distributions: Variations on a Theme of De Moivre. Statist. Sci. 6 (1991), pp. 284–302.
[11] Döbler, C., Gaunt, R. E. and Vollmer, S. J. An iterative technique for bounding derivatives of solutions of Stein equations. Electron. J. Probab. 22 no. 96 (2017), pp. 1–39.
[12] Döbler, C. and Peccati, G. The Gamma Stein equation and non-central de Jong theorems. Bernoulli 24 (2018), pp.384–3421.
[13] Eden, R. and Vieux, F. General upper and lower tail estimates using Malliavin calculus and Stein’s equations. In: Seminar on Stochastic Analysis, Random Fields and Applications VII, Eds: Dalang, R. C., Dozzi, M. and Russo, F., Progress in Probability 67, Birkhauser, 2013.
[14] Eden, R. and Vigueu, J. Nourdin-Peccati analysis on Wiener and Wiener-Poisson space for general distributions. Stoch. Proc. Appl. 125 (2015), 182–216.
[15] Eichelsbacher, P. and Thäle, C. Malliavin-Stein method for Variance-Gamma approximation on Wiener space. *Electron. J. Probab.* **20** no. 123 (2015), pp. 1–28.

[16] Gaunt, R. E. Variance-Gamma approximation via Stein’s Method. *Electron. J. Probab.* **19** no. 38 (2014), pp. 1–33.

[17] Gaunt, R. E. On Stein’s method for products of normal random variables and zero bias couplings. *Bernoulli* **23** (2017), pp. 3311–3345.

[18] Gaunt, R. E. A Stein characterisation of the generalized hyperbolic distribution. *ESAIM: PS* **21** (2017), pp. 303–316.

[19] Gaunt, R. E. Products of normal, beta and gamma random variables: Stein operators and distributional theory. *Braz. J. Probab. Stat.* **32** (2018), pp. 437–466.

[20] Gaunt, R. E., Mijoule, G. and Swan, Y. Stein operators for product distributions. [arXiv:1604.06819v2], 2016.

[21] Gaunt, R. E., Pickett, A. M. and Reinert, G. Chi-square approximation by Stein’s method with application to Pearson’s statistic. *Ann. Appl. Probab.* **27** (2017), pp. 720–756.

[22] Goldstein, L. and Reinert, G. Distributional transformations, orthogonal polynomials, and Stein characterisations. *J. Theoret. Probab.* **18** (2005), pp. 237–260.

[23] Götze, F. On the rate of convergence in the multivariate CLT. *Ann. Probab.* **19** (1991), pp. 724–739.

[24] Johnson, N. L., Kotz, S. and Balakrishnan, N. *Continuous Univariate Distributions*, Vol. 2. 2nd ed. Wiley, 1995.

[25] Kusuoka, S. and Tudor, C. A. Stein’s method for invariant measures of diffusions via Malliavin calculus. *Stoch. Proc. Appl.* **122** (2012), pp. 1627–1651.

[26] Ley, C., Reinert, G. and Swan, Y. Stein’s method for comparison of univariate distributions. *Probab. Surv.* **14** (2017), pp. 1–52.

[27] Luk, H. *Stein’s Method for the Gamma Distribution and Related Statistical Applications*. PhD thesis, University of Southern California, 1994.

[28] Nourdin, I. and Peccati, G. Stein’s method on Wiener chaos. *Probab. Theory Rel.* **145** (2009), pp. 75–118.

[29] Nourdin, I. and Peccati, G. Cumulants on the Wiener space. *J. Funct. Anal.* **258** (2010), pp. 3775–3791.

[30] Nourdin, I. and Peccati, G. *Normal Approximations with Malliavin Calculus: From Stein’s Method to Universality*. Cambridge Tracts in Mathematics. Cambridge University Press, 2012.

[31] Nourdin, I., Peccati, G. and Réveillac, A. Multivariate normal approximation using Stein’s method and Malliavin calculus. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** (2010), pp. 45–58.

[32] Nourdin, I. and Poly, G. Convergence in law in the second Wiener/Wigner chaos. *Electron. Commun. Probab.* **17**, no. 36, (2012), pp. 1-12.

[33] Olver, F. W. J., Lozier, D. W., Boisvert, R. F. and Clark, C. W. *NIST Handbook of Mathematical Functions*. Cambridge University Press, 2010.

[34] Peccati, G. Quantitative CLTs on a Gaussian space: a survey of recent developments. *ESAIM: Proceedings* **44** (2014), pp. 61–78.

[35] Peköz, E. and Röllin, A. New rates for exponential approximation and the theorems of Rényi and Yaglom. *Ann. Probab.* **39** (2011), pp. 587–608.

[36] Peköz, E., Röllin, A. and Ross, N. Degree asymptotics with rates for preferential attachment random graphs. *Ann. Appl. Probab.* **23** (2013), pp. 1188–1218.

[37] Pike, J. and Ren, H. Stein’s method and the Laplace distribution. *ALEA Lat. Am. J. Probab. Math. Stat.* **11** (2014), pp. 571–587.

[38] Stein, C. A bound for the error in the normal approximation to the the distribution of a sum of dependent random variables. In *Proc. Sixth Berkeley Symp. Math. Statis. Prob.* (1972), vol. 2, Univ. California Press, Berkeley, pp. 583–602.

[39] Stein, C. *Approximate Computation of Expectations*. IMS, Hayward, California, 1986.