Beyond Convexity: A Mathematical Summary and Derivation

Summary and Derivations Prepared by ChatGPT

Abstract

This document presents a complete, mathematically detailed summary of Jessica James' research note "Beyond Convexity" (Commerzbank, 2023). We derive the expressions behind bond duration, convexity, and the higher–order Taylor terms (third and fourth derivatives of price with respect to yield). These higher terms become essential for super-long bonds (50–100 year maturities), especially in low-yield environments (0–1%). We explain the behaviour of these expansions, the breakdown of convexity-based intuition, and the practical implications for hedging long-bond exposures using long-term swap rates. All figures from the original paper are omitted; instead, their qualitative information is described in the text.

Contents

1 Introduction

Traditionally, the sensitivity of a bond's price to yield changes is summarised using two quantities:

- Duration (first derivative of price with respect to yield),
- Convexity (second derivative).

For ordinary maturities (up to roughly 30 years), these two terms provide an excellent approximation of the price–yield relationship for moderate yield changes (e.g. $\pm$50–100bp). However, in the case of very long bonds (50y–100y), and in particular when yields are extremely low (0–1%), the price–yield curve becomes *so curved* that duration + convexity no longer give correct results.

This paper derives the full Taylor expansion of bond returns

$$R = \frac{\Delta PV}{PV}$$

and shows that the *third* and *fourth* derivatives become quantitatively important for 100-year bonds.

2 Present Value Representation

Let a fixed-coupon bond have:

- Coupon rate c (annualised, expressed as a decimal, e.g. $c = 0.02$),
- Yield to maturity y,
- Maturity n (in years),
- Unit notional.

The bond's present value is

$$PV(c, y, n) = \sum_{k=1}^{n} \frac{c}{(1+y)^k} + \frac{1+c}{(1+y)^n}. \tag{1}$$

This is a finite geometric series.

Define

$$V = \frac{1}{1+y}. \tag{2}$$

Then the PV can be rewritten in closed form:

$$PV = cV\frac{1-V^n}{1-V} + V^n. \tag{3}$$

Equation (3) is the key starting point for computing derivatives with respect to yield. It expresses the PV entirely in terms of the variable $V(y)$.

3 Bond Return and Taylor Expansion

Define the (relative) bond return for a yield move Δy:

$$R = \frac{\Delta PV}{PV}.$$

We expand $PV(y + \Delta y)$ in a Taylor series around y. The expansion for R becomes:

$$\begin{aligned} R = {} & \frac{1}{PV}\frac{dPV}{dy}\Delta y + \frac{1}{2}\frac{1}{PV}\frac{d^2PV}{dy^2}(\Delta y)^2 + \frac{1}{6}\frac{1}{PV}\frac{d^3PV}{dy^3}(\Delta y)^3 \\ & + \frac{1}{24}\frac{1}{PV}\frac{d^4PV}{dy^4}(\Delta y)^4 + \cdots \end{aligned} \tag{4}$$

The first term corresponds to *duration*, the second to *convexity*, and the remaining terms are the higher derivatives. For long-tenor, low-yield bonds these higher derivatives are not negligible.

4 Derivatives of PV with respect to Yield

All derivatives follow from the chain rule applied to $V(y) = (1+y)^{-1}$. We list the components:

$$\begin{aligned} \frac{dV}{dy} &= -(1+y)^{-2} = -V^2, \\ \frac{d^2V}{dy^2} &= 2(1+y)^{-3} = 2V^3, \\ \frac{d^3V}{dy^3} &= -6(1+y)^{-4} = -6V^4, \\ \frac{d^4V}{dy^4} &= 24(1+y)^{-5} = 24V^5. \end{aligned}$$

The PV formula (3) must be differentiated term-by-term. To keep the exposition readable, we record only the final derivative expressions, leaving intermediate algebraic manipulations aside. (These can also be automated in symbolic algebra systems.)

4.1 First Derivative (Duration Term)

We define *modified duration* D through:

$$D = -\frac{1}{PV}\frac{dPV}{dy}.$$

Differentiating (3), one obtains:

$$\frac{dPV}{dy} = -cV^2\frac{1-V^n}{1-V} - cV\frac{d}{dy}\left(\frac{1-V^n}{1-V}\right) - nV^{n+1}. \tag{5}$$

The middle term involves derivatives of quotient $\frac{1-V^n}{1-V}$ and introduces V^{n+k} terms. For long maturities, V^n is very small unless yields are very low.

4.2 Second Derivative (Convexity Term)

Convexity is

$$C = \frac{1}{PV}\frac{d^2PV}{dy^2}.$$

The second derivative follows by differentiating (5). The resulting expression contains combinations of V^2, V^3, V^{n+2}, V^{n+3}, etc. The magnitude increases dramatically as $y \to 0$.

4.3 Third Derivative

The third derivative is

$$T_3 = \frac{1}{PV}\frac{d^3PV}{dy^3}.$$

The expression is long but qualitatively follows:

$$\frac{d^3PV}{dy^3} = \alpha_1 V^4 + \alpha_2 V^5 + \alpha_3 n V^{n+3} + \alpha_4 n^2 V^{n+3} + \alpha_5 n^3 V^{n+3},$$

with coefficients α_i depending on coupon c. The key observation is that as $V = (1+y)^{-1}$ approaches 1, all V^k terms decay slowly, and the pre-factors involving n cause enormous magnification.

4.4 Fourth Derivative

Similarly,

$$T_4 = \frac{1}{PV}\frac{d^4PV}{dy^4},$$

and

$$\frac{d^4PV}{dy^4} = \beta_1 V^5 + \beta_2 V^6 + \beta_3 n V^{n+4} + \beta_4 n^2 V^{n+4} + \beta_5 n^3 V^{n+4} + \beta_6 n^4 V^{n+4}.$$

For a 100y bond at 1% yield, the term $n^4 V^{n+4}$ becomes large enough that $(\Delta y)^4$ contributions are comparable to the $(\Delta y)^2$ convexity term for moves of a few hundred basis points.

5 Asymptotic Behaviour as yield Approaches Zero

When yields approach zero,

$$V = \frac{1}{1+y} \approx 1 - y + y^2 - \cdots ,$$

and long-tenor cashflows become almost undiscounted.

For the derivatives:

$$\frac{dPV}{dy} = O\left(\frac{n}{y}\right), \qquad \frac{d^2PV}{dy^2} = O\left(\frac{n}{y^2}\right),$$

$$\frac{d^3PV}{dy^3} = O\left(\frac{n}{y^3}\right), \qquad \frac{d^4PV}{dy^4} = O\left(\frac{n}{y^4}\right).$$

Thus, the Taylor terms scale like:

$$(\Delta y)^k \cdot \frac{1}{y^k},$$

so if $y = 1\%$ and $\Delta y = 2\%$, the factors $(\Delta y/y)^k$ become huge (2, 4, 8, 16).

This explains why convexity (quadratic term) alone produces qualitatively wrong behaviour for 100y bonds in a low-yield environment.

6 Why Convexity Fails for Super-Long Bonds

The convexity contribution to return is

$$\frac{1}{2}C(\Delta y)^2,$$

which is *always positive*. Thus, for sufficiently large $\Delta y > 0$, the quadratic term dominates the linear duration term and predicts that price increases with yield, which is impossible (bond value must fall).

In the figures of the original paper (described):

- For a **30y bond**, the true return is well approximated by duration + convexity for $\pm 1\%$ yield change.
- For a **100y bond**, convexity alone bends the approximation upwards for large positive Δy, giving the illusion that bond price may increase for sufficiently large rate rises.
- The **third term** (cubic) corrects this behaviour, pulling the curve back down for positive Δy, and upward for negative Δy (asymmetric shape).
- The **fourth term** (quartic) makes the approximation match the true return almost perfectly even for $\pm 2\%$ shifts.

7 Interpretation of Figures (Textual Explanation)

Figure: Price–Yield Return Curves (30y vs. 100y)

The original chart shows that:

- 30y bonds have a smooth, modestly curved price–yield relationship.
- 100y bonds have an extremely curved relation—returns are much more sensitive to even small yield changes.

Figure: Taylor Approximations vs. True Return

- For 30y bonds: duration + convexity is nearly exact.
- For 100y bonds:
  - Duration only: huge error.
  - Duration + convexity: wrong curvature for large Δy.
  - Add third term: major improvement.
  - Add fourth term: nearly perfect.

Residuals Plots

Residual = true return minus approximation.

- Convexity produces quadratic-shaped residuals.
- Third term corrects one side but over-corrects the other.
- Fourth term removes almost all remaining structure.

8 Real-World Case Studies

Two important examples in the paper:

- **NRW 2119** (100-year bond)
- **OAT 2072** (50-year French government bond)

For investors who bought these bonds at yield lows (2020–2021):

- The NRW 100y bond suffered losses of up to **70%** after a 250bp rise.
- Duration alone drastically underestimates losses.
- Duration + convexity still gives incorrect results.
- Only the inclusion of third and fourth Taylor terms can replicate the observed returns.

9 Hedging 100-Year Bonds with 50-Year Swaps

The paper studies a hedge that is:

$$\text{Hedge ratio} = \frac{\text{Duration of bond}}{\text{Duration of swap}}.$$

For the NRW bond:

$$D_{\text{bond}} \approx 58.8, \qquad D_{\text{swap}} \approx 55.7,$$

yielding a hedge ratio of approximately 1.06.

The result:

- Duration terms cancel by construction.
- The hedge leaves the investor long convexity: positive returns for both rising and falling yields.
- Third and fourth terms reduce this benefit for large rate rises, and amplify it for rate falls.

Qualitatively, the hedged 100y bond still had positive performance during the stress period due to the long-convexity position.

10 Conclusions

- Traditional duration + convexity intuition breaks down for super-long bonds (50y–100y) at low yields.
- The Taylor expansion of bond returns requires at least up to the **fourth** derivative for accurate results.
- Convexity alone eventually predicts the wrong sign for returns under large positive rate moves.
- Third and fourth derivatives correct this behaviour.
- Hedging long bonds with long-term swaps produces positive convexity exposure, but requires awareness of higher-order risks.